Measurement of the inclusive-jet cross-section in proton-proton collisions and study of Quark-Gluon Jet discrimination with the ATLAS experiment at the LHC

HILAL KUCUK

Submitted to University College London in fulfillment
of the requirements for the award of the
degree of Doctor of Philosophy

University College London
April 12, 2016

Declaration

I, Hilal Kucuk, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Hilal Kucuk 

Abstract

Several jet measurements have been performed using data collected during the 2011 and 2012 data taking periods at respectively 7 and 8 TeV proton-proton collision energy by the general-purpose ATLAS experiment at the LHC. These measured are compared to Leading Order and Next Leading Order Monte Carlo predictions. This thesis works focalises in particular on two topics: the measurement of the inclusive double-differential jet cross section, and the separation between jets originated by quarks and those originated by gluons, motivated by its possible use in searches for contact interactions in dijet events.

In the first part of thesis, the inclusive double-differential cross section has been measured at $\sqrt{s}$=8 TeV on low-luminosity runs taken in 2012, as a function of jet transverse momentum in bins of jet rapidity, covering a range of $25 < p_T < 134$ GeV and $|y| < 3.0$. Jets are defined using the anti-k_t algorithm with two radius parameters of 0.4 and 0.6. The data sample has been collected using two special runs: a no-pile up run, and a low mu run using data-taking periods from A2 an A6. The total luminosity of the no pile-up and the low μ runs is 193.3 nb^{-1}. The reason for using these two runs is to measure the low-pT region for jets, without the very large systematic uncertainties that would be caused by pileup in this kinematic region. This measurement has been merged with the high mu run cross-section measured by Gagik Vardanyan, for the ATLAS publication on the inclusive double-differential cross section measurement at $\sqrt{s}$=8 TeV.

Separating quark-initiated jets from gluon-initiated jets can provide improve the reach of many beyond-the-standard model searches such as compositeness and contact interactions. In order to distinguish quark from gluon jets, a likelihood was built based on the number of charged track in the jet, and on the track width for the analysis of the 2011 dataset. Once jets have been

labelled as quark-originated or gluon-originated, dijet events can be labelled as quark-quark, quark-gluon, and gluon-gluon. The detector-level invariant mass spectrum has been measured in bins of y^* in 2011 data, and compared to Standard Model QCD or models of contact interactions with different Λ scales. The tool has been further improved to have a better quark-gluon jet discrimination for the 2012 analysis. Calorimeter-based variables have been added to the track-based ones used in 2011, to extend the rapidity coverage and the discrimination power of the tool. The additional advantage of using 2012 data is the increased luminosity, from the 4.7 $\pm$0.2 fb^{-1} of 2011 to the 20.3 fb^{-1} for 2012, allowing the use of data-driven techniques to measure the jet shapes used in the discrimination. As quark and gluon fractions and jet properties are strongly dependent on jet kinematics, all variables are studied in bins of p_T and η.

ACKNOWLEDGMENTS

I would like to express my deep gratitude to everyone who has helped towards this work, especially Shima Shimizu, Christopher Meyer,Brinick Simmons. Thanks a lot to my supervisor Mario Campanelli for both his support and for his unerring patience to me. I also need to appreciate his very quick answer via e-mail like speed of light. Thanks also to Pauline Bernat and Alex Christopher Martyniuk for their help, smile and patience with my many naive questions. I do not know how Mario, Pauline and Alex bear with me. Thank you to Brinick Simmons, for helping make sense of the complex ATLAS software testing systems.

Thank you also to Stephen Jiggins and Valentinos Christodoulou to make CERN office silent. How I can forget my best office mates who Peter Davison and Thomas Stuttard and their football conversations.

TURKISH ACKNOWLEDGMENTS

Bu tezi siz aileme armagan etmek istiyorum. Ingilizce,master, doktora derken 7 koca yil sizsiz gecirdim. Sizden ayri kalmamak icin arabayla 1 saat 15 dakika uzakliktaki Kirikkale'den bile universite tercihi yapmamisken, kader beni ucakla 7-8 saat suren Ingiltere'ye yolladi hem de sizsiz. Cok zordu ama sizin desteginizle oldu. O kadar sizsiz kaldim ki cok agir hastalandigimda, su anda son nefesimi versem annemgilin bana ulasmasi 1-2 gun surecek diye dusundum cogu kez. Fiziksel olarak olmasanizda, varliginiz ve dualariniz olmasaydi asla bu okulu bitiremezdim. Hele annem ve ablamin dualari, onlarla neler atlattim yazsam bitmez. Benim uc annem var diyorum, biri bana durust olmayi, ahlakli olmayi, dini ogreten annem biri de anne yurekli, hakkini asla odeyemiyecegim biricik babam, nasil guzel bir babasin sen ve ben zora dustugumde kendi sorunlarini unutup bana kosan benim icin aglayan dunyalar biricigi ablam. Olmasaydiniz nasil biri olurdum acaba. Tabi yigenlerim ve ikizim olan abimi unuturmuyum. Ikizim diyorum, ruh ikizim aslinda ikizim degil. Annem ikiniz de ruh olarak aynisiniz, yalnizca sen tembellikten 7 yil karnimda yatmissin, gec dogmussun diyordu. Iyi ki varsin abim benim. Ve tabii ki Demet ve ikinci abim Yavuz abiyi unutmus degilim. Yigenlerime gelince, Merve,Mert, Eren. Buraya iyi ki gelmisim desem de hep sizden dolayi bir yanim buruk oldu. Cunku sizler buyurken ben sizinle olamadim. Ilk askinizda, ilk yere dususunuzde yaninizda bulunamadim. Herkese yanimda olduguklari icin tesekkur ediyorum ama sizden yaninizda olamadigim icin ozur diliyorum, keske sizin uzaginizda ki teyzeniz/halaniz olmasaydim da hep yaninizda olabilseydim. Paraniz bittiginde ilk geleceginiz kisi ben olsaydim. Ama sizler benim en sikintili animda, yaptiginiz sirinliklerden dolayi yasama sevincim oldunuz. Sizleri yigenim degil, cocuklarim olarak goruyorum. Iyi ki varsiniz. Bu yuzden bu tezi illa ki size armagan ediyorum.

Ve Londrada rastladigim, 7 yil suresinde bana dost olan, nese veren arkadaslarima cok tesekkur etmek istiyorum. Ama Rumeysa Bayar'a hem tesekkur hem de

ozur borcluyum. Hem daim nazimi ve kahrimi cektigin icin ve beni surekli simarttigin icin cok tesekkur ederim. Tabii herkese benim icin zorla veya isteterek dua ettirdigin icin de extra tesekkuru borc bilirim. Tezin bu kisminda yalnizca tesekkur edilir ama mecburen kendisini istemeden kirdigim icin de ozur dilerim insallah kabul eder.

Ilk tanistiklarimdan basliyarak devam edecegim. Nevin Meral ve Adem aslinda size tesekkur etmeli miyim bilmiyorum, iliskiniz sayesinde en az 6-7 ay doktorami uzattim desem yalan olmaz. Tabii ki saka, iyi ki vardiniz. Merve Altinli, sohbetlerin, bir sey gorunce heycanlanmalarin bana cok sey katti, sana da tesekkur ederim. Husrev Tabak her ne kadar arada sirada fikir catismalari yasasakta, Islamci abilere bakis acimi degistirdin sagolasin, seninle sohbetler cok guzeldi. Tugba Cinar seninle kardes olmak guzeldi. Ve Aybike Salman, Deniz Ozkulu, Sultan Ileri hayatima guzel seyler kattiniz. Tabii Sultan daha cok catisma katmis olsa da :)

Zeynep Arslan hanim size de iki cift sozum var. Ne diyeyim bana Londrada yaptigin arkadasliklardan dolayi cok tesekkur ederim, tabii Londradaki restoranlarda tesekkur etmeli, neredeyse hepsini gezdik sayilir. Zeliha Kaplan sana da tesekkur mu etsem etmesem bilemedim, doktoramdam 3-4 ay calmisligin vardir :) Ayse Guven hanim size nasil tesekkur edilir bilmem, insan buyuklerinin tecrubelerinden yararlanir genellikle ama ben kucugum olan senden cok sey ogrendim. Ne diyeyim mubarek ve guzel insan, Londra'da ve dunyada tanimasaydim obur tarafa bos giderdim dedigim insanlardansiniz. Tabii Feyza Tunal sende bu guzel insanlara dahilsin. Huriye Atilgan, vallahi insanlar genelde isimleri gibi olurmus sen soyisminin hakkini veriyorsun. Karsilastigim zor ama sevdigim karakterlerdensin. Her ilmi konusmamizda teke tekken cok iyi anlasip, baskalarinin yanindayken tartistigim guzel insan sana da tesekkurler. Tabii ki shadwell ve kuzeydeki guzel insanlar size de ayrica tesekkurler. Ramazan Gundogdu ve Osman Ersoy'a da tesekkuru borc bilirim. Cok iyi bir kardes olduklari icin ve her zor durumda imdadima yetistikleri icin. Tabii Ilhan Candan hocamin kardesligine de cok tesekkur ederim,

sinirlendigimde kahve molalari icin de.

Contents

List of Figures

List of Tables

Chapter 1

Theoretical Framework

'In the name of God, the infinitely Compassionate and Merciful.' -The Opening in Quran.

1.1 Introduction

The Standard Model (SM) is a theory of fundamental interactions, and describes the behaviour of fundamental particles like quarks, leptons and bosons [1] as described in section 1.2. The elementary fermions and bosons of the SM with their masses, spins, charges and interactions between each other are shown in Figure 1.1. Electroweak interaction are discussed in section 1.4 in this chapter. The interactions described by QED and QCD (section 1.3)are carried by massless spin-1 particles, called gauge bosons(photon in QED, gluons in QCD) [2].

In this chapter, there is some discussion about some basic QCD phenomenology such as Evolution of the strong coupling constant, color confinement and Asymptotic freedom in subsection 1.5.1, also Parton Density Functions (PDFs) in subsection 1.5.2, QCD factorisation in subsection 1.5.3. Various methods of defining jets (subsection 1.5.4) that are commonly used by the ATLAS experiment and for the study in this thesis. Additionally, it also mentions the research for new physics (section 1.6), and a theory called "contact interaction" (subsection 1.6.1) which studied in this thesis.

1.2 The Standard Model

The W and Z bosons were discovered in 1983, the top quark in 1995, and the tau neutrino in 2000 [3]. A new particle which is consistent with the Higgs boson in the mass region around 126 GeV has been observed in 2012 by the ATLAS and CMS experiments at CERN's Large Hadron Collider [4]. The research is still under way to measure this particle's properties, and determine whether or not this particle is the Higgs boson predicted by the minimal Standard Model or by alternative models.

The SM is based on the SU(3)×SU(2)×U(1) symmetry groups, corresponding to strong, weak and electromagnetic interactions, respectively. Because of the symmetries of the theory and of Noether's theorem, quantities such as energy, angular momentum and electric charge must be conserved [5].

Atoms and molecules are held together through the electromagnetic force which is carried mainly by the photon. The weak force is mediated by the W^-,W^+ and Z^0 gauge bosons. It is responsible for several processes like the nuclear β decays, or the decay of the muon.

Particles decay into lighter states, conserving energy and momentum. The process binding quarks and gluons into hadrons which is groups consisted of quarks, and nucleons into an atomic nucleus is governed by the strong interaction. Strong interaction is stronger than electromagnetism, which is stronger than the weak nuclear force and gravity. The gravitational force is not included in the Standard Model, but is described in classical language by the general theory of relativity. Conserved quantities are found by exploiting the symmetries of the SM Lagrangian, and the equations of the separate electromagnetic, weak and strong theories can be derived from it.

Historically, the first component of standard model to be developed is quantum electrodynamics, that has been unified with Fermi's theory of weak interaction in the electroweak theory [7]. The description of strong interactions is formally quite similar to electroweak ones with some differences that will be

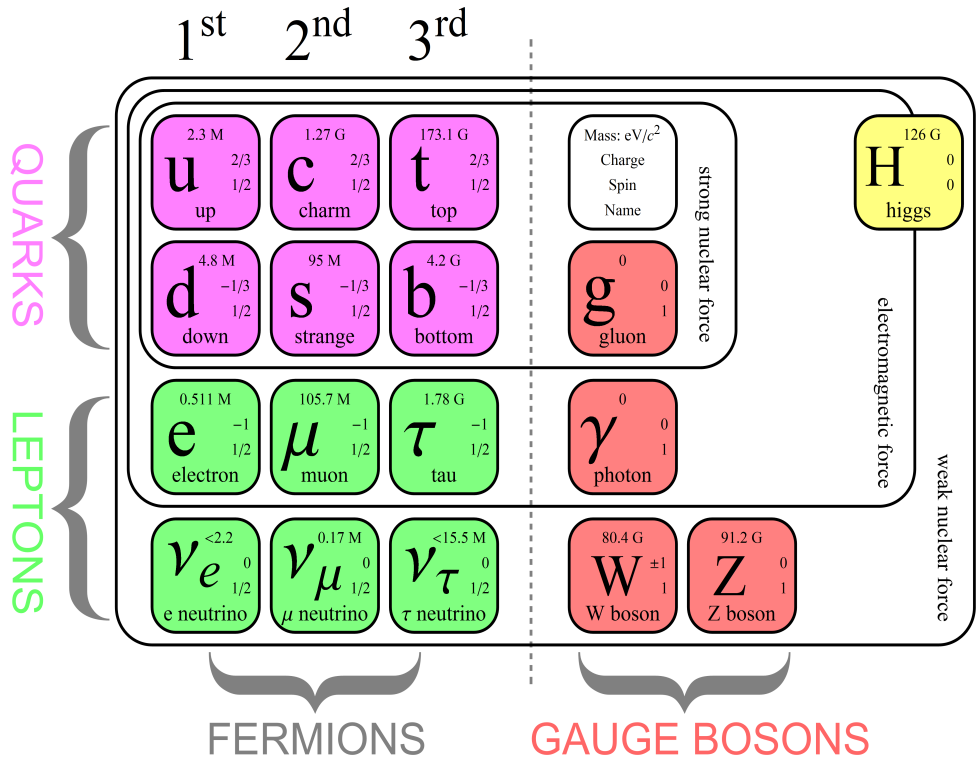


Figure 1.1: The Standard Model with masses, spin, charges of particles. The interaction of particles between each other is also shown [6].

examined with discussion at following section.

1.3 QED and QCD

In QED, electrons and positrons as charged particles interact by absorbing and emitting virtual photons. The virtual particles cannot be observed outside of the interaction because their life-time is incredible short [8]. The QED Lagrangian is defined as

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi - e\bar{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.1}$$

where $F_{\mu\nu}$, is the electromagnetic field tensor, and is defined, for a photon field A_{μ}, as

$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} \quad and \quad \bar{\psi} \equiv \psi^{\dagger}\gamma^{0} \tag{1.2}$$

where e is the electric charge, the γ^μ are the Dirac matrices and ψ is a bispinor field of spin-1/2 particles.

Quarks form the fundamental representation of the $SU(3)$ color group. As this representation is three-dimensional, quarks exist in three different colors, called red, green and blue for particles (anti-colors for antiparticles). Gluon can form a color octet because N^2-1 for N=3 is equal to 8. The QCD Lagrangian is defined as

$$\mathcal{L}_{QCD} = \sum_{f=1}^{n_f} \bar{\psi}_f (i\gamma^\mu D_\mu - m_f)\psi_f - \frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a \tag{1.3}$$

where the covariant derivative is given by

$$D_\mu \psi_f = \partial_\mu \psi_f + i g_s A_{\mu a} \frac{1}{2} \lambda_a \psi_f \tag{1.4}$$

and the field strength tensor for the gluon field, $F^{\mu\nu}_a$,is

$$F^{\mu\nu}_a = \partial^\mu A^\nu_a - \partial^\nu A^\mu_a - g_s f_{abc} A^\mu_b A^\nu_c \tag{1.5}$$

where f_{abc} are the structure constants of SU(3), and the a,b,c indices are for the eight color degrees of freedom of the gluon field. In Equation 1.3, the ψ_f are quark-field spinors for a quark of flavour f and mass m_f. The sum over f is for the different quark flavours: up,down, charm,strange, top and bottom. λ_a are the hermitian, traceless Gel-Mann matrices which allow for the rotation in color space and $A_{\mu a}$ are the spin-1 gluon field. g_s is gauge coupling parameter [9].

Since equations 1.3 contains linear and quadratic terms of the gluon field, their product can lead to couplings between 3 or 4 gluons resulting in gluon self interactions showed in figures 1.2.a and 1.2.b. Figure 1.2.c shows typical interactions between quarks mediated gluon exchange. Due to conserve electric charge, quark flavour is not changed by gluon emission even if the color can change.

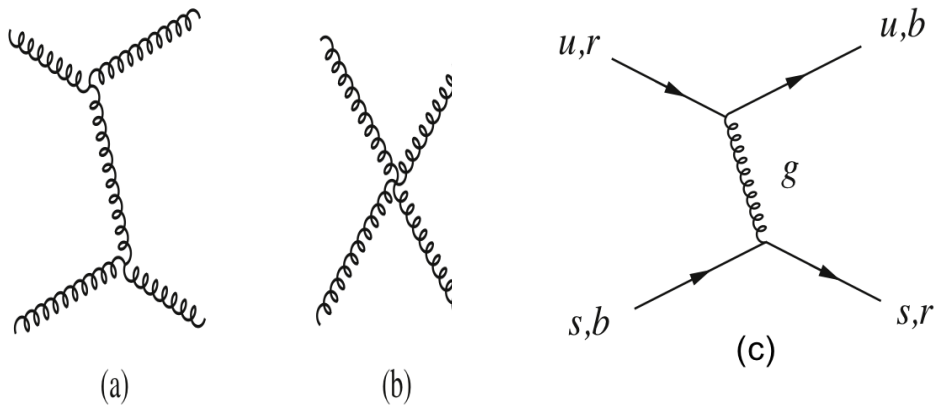


Figure 1.2: (a) and (b): the two lowest-order contributions to gluon-gluon scattering in QCD. (c): example of quark and quark scattering by gluon exchange [10].

1.4 Electroweak Interactions

The weak nuclear force has been initially described using the Fermi theory, that postulated the existence of a vertex with four quark or lepton lines. After the development of QED, it was understood that this vertex was actually the result of the exchange of a heavy vector boson. An initial theory describing these particles, now known as the $W^{\pm}$ bosons, was formulated. Similarities between the structure of electromagnetic and weak interactions have lead S. Glashow, in the early sixties, to formulate a unified theory of the two, through a Lagrangian symmetric under the SU(2)×U(1) gauge group. This approach predicted the existence of a neutral vector boson, the Z^0, mixing with the photon, and therefore the existence of weak neutral currents. The theory has been completed a few years lated by Weinberg and Salam, that included a mechanism for accounting for the different masses of the vector bosons and of the fermions through spontaneous symmetry breaking, hence the introduction of a new field and a new particle, the Higgs boson [11]. An example of a process mediated by the weak interaction can be seen in Figure 1.3, where two quarks, or a quark-antiquark pair scatter through the exchange of charged W

or neutral Z bosons.

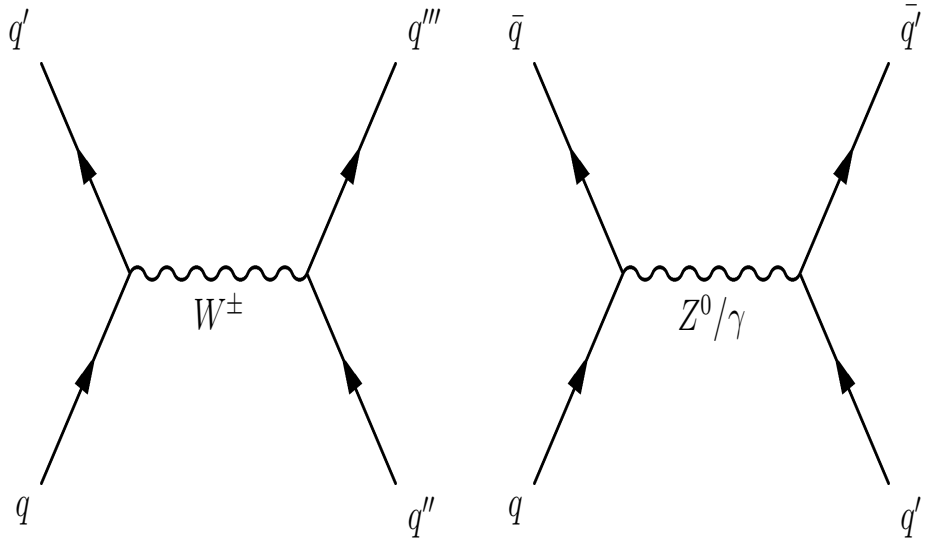


Figure 1.3: The decays of W,Z and γ bosons [12].

1.5 Some basic QCD phenomenology

1.5.1 Evolution of the strong coupling constant, color confinement and Asymptotic freedom

While for electromagnetic interactions the coupling constant α_{em} increases with momentum transfer, the non-Abelian structure of strong interactions (basically the fact gluons interact each other) has as a consequence the fact that the coupling constant decreases with Q^2 as shown in equation;

$$\alpha_s(Q^2) \simeq \frac{\alpha_s(Q_0^2)}{1 + B\alpha_s(Q_0^2) ln(\frac{\alpha_s(Q^2)}{\alpha_s(Q_0^2)})} \tag{1.6}$$

where the constant B is a function of the number of quark flavours, N_f, and of colors N_c:

$$B = \frac{11N_c - 2N_f}{12\pi} \tag{1.7}$$

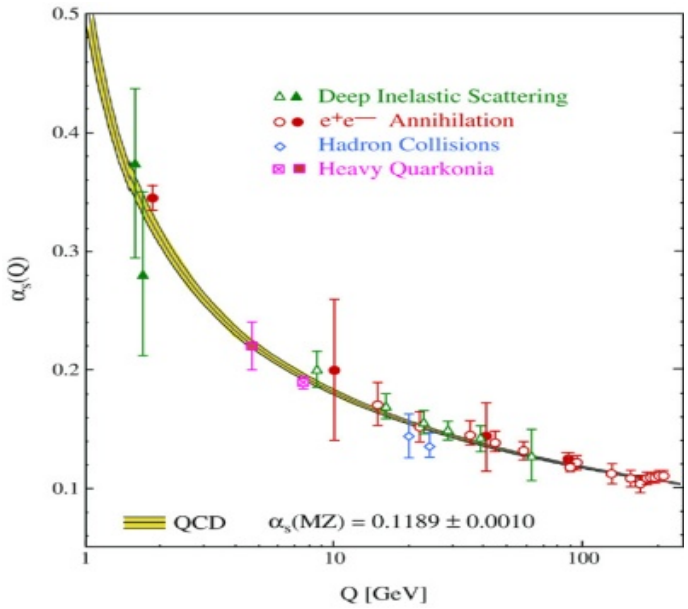


Figure 1.4: Coupling constant of the Strong Interaction wrt Exchanged Momentum [13].

The momentum Q_0 corresponds to scale $\Lambda_{QCD} \simeq 300$ MeV, that can be translated into distances of the order of 0.1 fm and is conventionally taken as the boundary between the perturbative and non-perturbative regime. Figure 1.4 shows some experimental measurements of the strong coupling constant as a function of the momentum transfer. Its decrease is clearly visible. Quarks and gluons inside a hadron are bound by strong interactions at low values of Q^2, where the coupling constant is very strong. The higher orders of perturbation theory become dominant, and no perturbative calculation will be able to give a good description of this process. If as a result of a collision one of the partons receives a large momentum boost, the distance between partons increases, creating a flow of color charge between them. The energy of this "color tube" will increase with distance, until it is sufficient to produce a new $q\bar{q}$ pair. This process continues until the color string has enough energy to produce new pairs. According to a property known as color confinement, all observable free particles must be colorless. It means that the colored particles such as quark and gluon have never been directly observed in the nature. The partons resulting from the process described above have therefore to confine in colorless

hadrons, which are observable. More details about hadronisation will be given later [14].

When low-momentum partons are bound inside the hadrons, strong interactions are weak at short distances (corresponding to large momentum transfer), so partons inside the hadrons can be considered as free. This is known as Asymptotic freedom [15]. If a high momentum scattering occurs, the lowest-order diagrams like the ones shown in figure 1.2 can be a reasonable approximation of the interaction. To date, Next-to-Leading Order (NLO) calculations are available for most of the processes of interest, and for a growing number of processes Next-to-Next-to-Leading Logarithmic approximations (NNL), or even full next-to-Next-to Leading Order (NNLO) calculations have been performed.

1.5.2 Parton Density Functions (PDFs)

The structure of hadrons in the regime of non-perturbative QCD is described by Parton Density Functions (PDFs), that represent the probability density of a parton inside the hadron, as a function of its fractional momentum with respect to that of the hadron, x, and the momentum transfer scale Q^2. Figure 1.5 shows the proton PDFs as a function of x for two different values of Q^2. It can be observed that the fraction of momentum carried by the high-x, "valence" quarks is higher at low Q^2, while the low-x part, corresponding to gluons and "sea" quarks", increases at higher Q^2 [16].

Since partons inside a proton are in the non-perturbative regime of QCD, it is not possible to calculate PDFs from first principles. For a given functional form, depending on various models, its parameters are derived from a global fits of a large number of experimental observables, from experiments of deep-inelastic scattering (with the HERA collider and fixed target experiments, also using muon and neutrino beams) and colliders like the Tevatron and the LHC. These experiments measure the same PDFs in a wide range of momentum transfers; it is possible to propagate these functions from a scale

to another using evolution equations, like DGLAP [17] or BFKL [18]. Various PDF fits are currently available, differing by the choice of the functional form, the dataset considered and the statistical treatment of the data. Through the recent precise LHC data, PDFs can be constrained studying inclusive jet productions.

The choice of PDFs is very important in the simulation of both hard processes and parton showers which is cascade of partons, since it affects production cross-sections and event shapes. Typical sets used at the LHC are CT10nlo [19], HERAPDF15NLO_EIG [20], NNPDF 3.0 [21] and MSTW2008-nlo68cl [22] in this thesis for the measurement of inclusive cross-section.

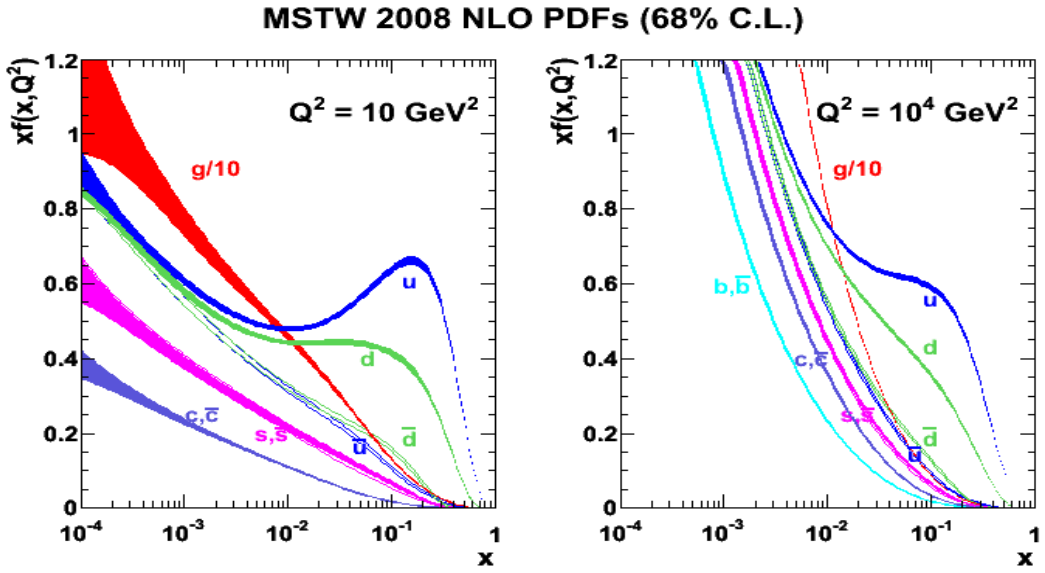


Figure 1.5: Parton distribution functions of the proton as determined for the MSTW08 PDF (Martin-Stirling-Thorne-Watt Parton Distribution Functions) set for (left) $Q^2 = 10\ GeV^2$ and (right) $Q^2 = 10^4\ GeV^2$. The bands reflect the uncertainties at the 68% confidence level [23].

1.5.3 QCD factorisation

A full calculation of a QCD process involves terms in both the perturbative and the non-perturbative regime, so several steps are needed. For the hard scattering process, perturbative calculations of the parton-parton scattering

are available, often at higher orders. Since the collision does not occur directly between partons but between protons, the hard scattering calculation has to be convoluted with the PDFs, that are the result of an empirical model. The final-state partons from the hard scattering will undergo a process called parton shower, where gluons are emitted and quark-antiquark pairs are created. At the end of this process, low-energy partons will be merged to produce hadrons, in a non-perturbative process called hadronisation. The combination of perturbative and non-perturbative techniques is made possible by the factorisation theorem, stating that the two regimes can be treated independently and then combined [24].

The application of QCD factorisation is formalised in Equation 1.8

$$\sigma_{jet} = \sum_a \sum_b \overbrace{f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2)}^{\text{PDFs}} \overbrace{\hat{\sigma}_{a,b}(p_{p1}, p_{p2}, \alpha_s(\mu_R^2), \frac{Q^2}{\mu_F^2}, \frac{Q^2}{\mu_R^2})}^{\text{Hard scatter}} \tag{1.8}$$

where μ_F used in the evaluation of PDFs and μ_R are factorisation scale and renormalisation scale respectively, f_a and f_b are the PDFs of the interacting protons. Typically, μ_R and μ_F are set equal to Q^2, but also varied by a factor of 2 to determine systematic uncertainties connected to the scale choice.

1.5.4 Hadronisation and Jets

It was already mentioned that at the end of parton showering, colored partons merge into colorless hadrons in a process called hadronisation.

Since it involves low momentum transfers, hadronisation is a non-perturbative process, and has to rely on empirical models. The two main hadronization models are cluster hadronization (Figure 1.6) and string hadronization (Figure 1.7) [26] [27]. The string model is based on a linear confinement at large distance, most easily simulated for the production of a heavy quark-antiquark pairs. A string is stretched from the quark to anti-quark via gluon which behaves like “kink“. An alternative model is the cluster model, where partons

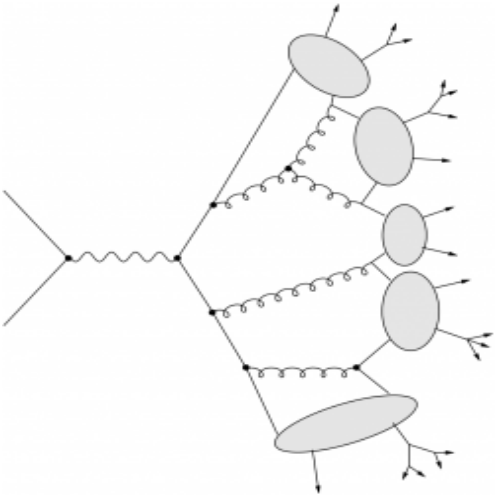


Figure 1.6: Cluster hadronization model [25].

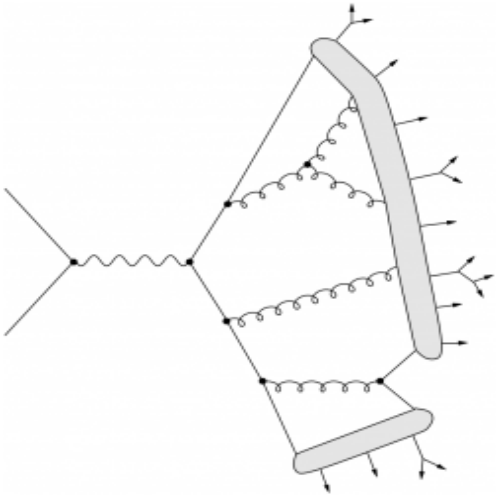


Figure 1.7: String hadronization model.

are grouped into colorless "clusters", that will eventually lead to hadrons.

The observable result of a high-momentum parton is a collimated "spray" of particles, called jet, emitted approximately in the same direction as the initial parton. Jets can be seen as the experimental proxy for production of high-energy quarks and gluons. Jets are the most common objets at hadron colliders, and the most common process producing them is $2 \rightarrow 2$ QCD scattering of partons. The jet measurements play an important role on searches for the standard model and physics beyond the standard model, on the measure of the structure of the proton and of QCD properties like its coupling strength α_s.

1.5.5 Jet algorithms

Final-state hadrons are clustered into jets using algorithmic jet definitions. Algorithms define which hadrons belong to which jets. While in the Monte Carlo simulation the final state hadrons are directly available from the event record, in data hadrons are reconstructed merging clusters from the energy depositions of electromagnetic showers (EM) [1] and hadronic showers in the

[1] EM showers are generated by a particle that interacts the electromagnetic force, usually a photon or electron

calorimeters. A proper jet algorithm must fulfil some requirements. It must be well-defined at any order of perturbation theory, and be both collinear and infrared safe, being stable under minor adjustments of the jet constituents.

Infrared safety means that the result of jet clustering should remain unchanged when adding a soft parton. Collinear safety refers that when replacing a single parton by a collinear pair of partons, the identified jets should not change [28].

Jet algorithm can be classified in two types: cone and recombination.

Cone algorithms have an apparent conceptual simplicity and a fast execution time. Jet reconstruction starts from a high-momentum cluster above a given threshold, taken as a seed. All clusters inside a given cone around the seed are considered to be part of the jet. Jets can also be merged and split according to some rules. Since soft radiation can affect the choice of the seed, cone algorithms are generally infrared and collinear unsafe. For this reason, both ATLAS and CMS moved from cone-based to recombination algorithms for jet reconstruction [29].

Recombination algorithms combine particles or calorimeter clusters according to quantities such as their p_T and direction. The processing for this type algorithms starts from the "closest" particles by defining a distance (d_{ij}) between them. Particles with the smallest distance are merged, and replaced by the result of the merging; this process continues iteratively until there is no object left. The distance used by most recombination algorithms is defined as

$$d_{ij} = min(k_{Ti}^{2p}, k_{Tj}^{2p})\frac{\Delta_{ij}^2}{R^2}, \quad d_{iB} = k_{Ti}^{2p} \tag{1.9}$$

where R is the jet radius parameter, $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, d_{ij} is the distance between i particle and j particle and d_{iB} represents distance between the beam direction and particle i. k_{Ti}, y_i and ϕ_i are respectively the transverse momentum, rapidity and azimuth of particle of i. The value of p defines the

type of recombination algorithm,

$$p = \begin{cases} 1, & k_\perp \text{ algorithm} \\ 0, & \text{Cambridge/Aachen algorithm} \\ -1, & \text{anti-}k_\perp \text{ algorithm} \end{cases} \tag{1.10}$$

The p_T ordering of the jet constituents is different between the clusterings. The simplest recombination algorithm is the Cambridge/Aachen algorithm, that depends on only angular quantities, because p=0. This dependency gives opportunity to study a range of angular scales when reconstructing the substructure of a jet.

The $k_\perp$ algorithm produces irregularly shaped jets since it combines clusters from pairs of low p_T objects. It is quite sensitive to the effects of pileup because of this.

The anti-k_t algorithm merges clusters starting from the high-p_T pairs, up to a distance corresponding to the jet radius parameter, producing circular jets. Using the anti-k_t algorithm allows to reduce sensitivity to pileup and make easier calibration. The other advantage is that, although almost circular jets are produced, the algorithm is still collinear and infrared safe [28].

In Figure 1.8, Cambridge-Aachen, anti-$k_\perp$ and $k_\perp$ are applied to clusters from the same event, shown in $\eta - \phi$ space. Similar directions for hard jets are reconstructed by all algorithms, but the attribution of the soft components, as well as the low-energy jets are different.

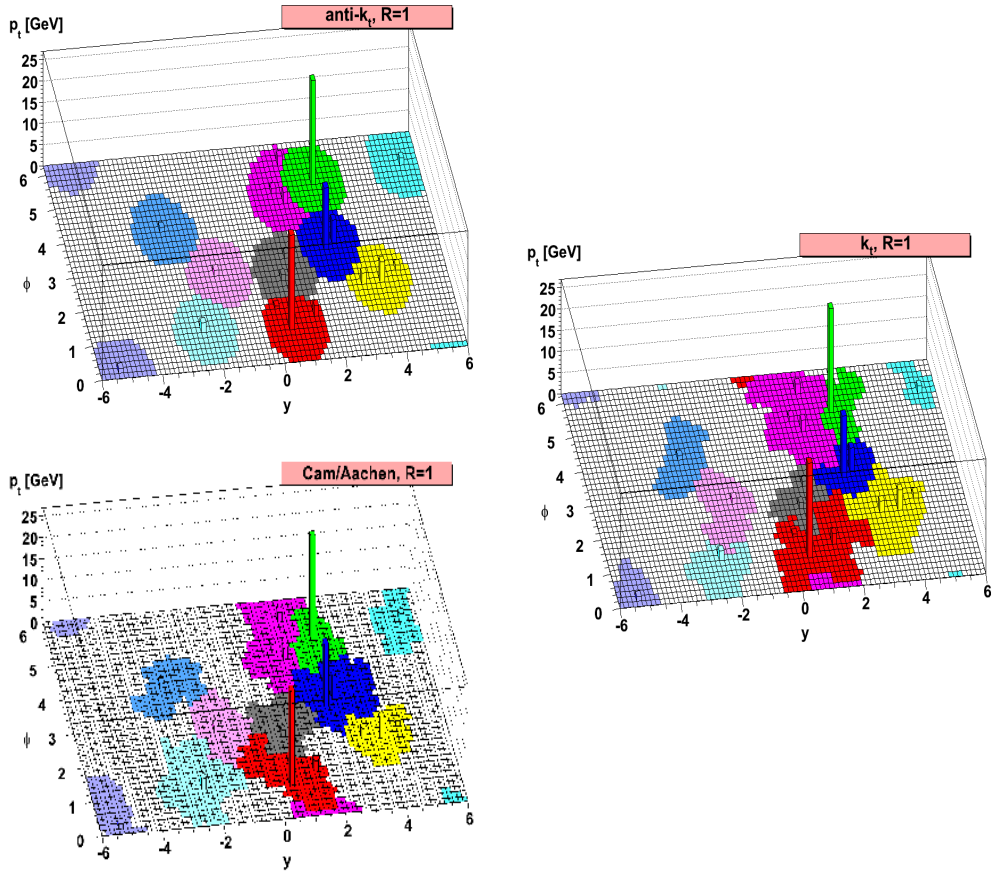


Figure 1.8: The same events are clustered using different jet clustering algorithms, anti-$k_\perp$ (top left), Cambridge-Aachen (right), $k_\perp$ (bottom) and in $\eta - \phi$ space [28].

1.6 The search for new Physics

Despite its success, the Standard Model is an incomplete theory: it does not account for gravity, and the strong force is "added" on top of the electroweak one, but not really unified. Moreover, if new particles exist in the range between the electroweak scale and the Plank scale, the Higgs mass will receive contributions from higher-order corrections, and its small value could only be explained if some basic parameters are unnaturally fine-tuned to a large precision. During the last decades, many new physics theories have been proposed to solve some of the shortcomings of the Standard Model, or to explain additional phenomena like dark matter or the baryon asymmetry of the uni-

verse [30]. Examples of such theories are Supersymmetry [31], new dynamics modelled on the strong force, the existence of extra space-time dimensions [32], etc. Signatures for these theories include excesses in events with high invariant masses of final-state particles, missing transverse momenta, isolated high p_T leptons, very energetic photons and jets. To extend the search of new physics beyond the current limits, higher energies and luminosities are needed, and even in the case of a positive signal, the data from the LHC may not be enough to separate different models.

1.6.1 Contact Interactions

A possible signature of the presence of new physics may be indicated by contact interactions. This term refers to an interaction mediated by a particle whose mass is much higher than the energy scale of available in the collision, but that would manifest itself in a modification of the cross-section for high invariant masses. The scale of new physics is indicated by the value Λ [33] [34]. A quark-quark scattering process can be modified by this additional interaction, leading to a modification of the final-state invariant mass distribution. Dijet production at high p_T for SM QCD is dominated by t-channel gluon exchange. Figure 1.9 shows two additional processes, where a new heavy particle of mass M, similar to the scale of new physics, Λ, is exchanged either in s-channel or t-channel. For values of this mass much larger than the dijet invariant mass, the two diagrams can be approximated by a single four-legs contact interaction, indicated by the diagram at the bottom. The traditional parametrisation of the interaction Lagrangian is:

$$\begin{aligned}\mathcal{L} = \frac{2\pi}{\Lambda^2}[&\eta_{LL}(\bar{q}_L\gamma^\mu q_L)(\bar{q}_L\gamma_\mu q_L)+ \\ &\eta_{RR}(\bar{q}_R\gamma^\mu q_R)(\bar{q}_R\gamma_\mu q_R)+ \\ &2\eta_{RL}(\bar{q}_R\gamma^\mu q_R)(\bar{q}_L\gamma_\mu q_L)\end{aligned} \tag{1.11}$$

where the parameter η_{ij} (i and j can take the values Left or Right), de-

fines the coupling of this new interaction to either left-handed or right-handed quarks [35].

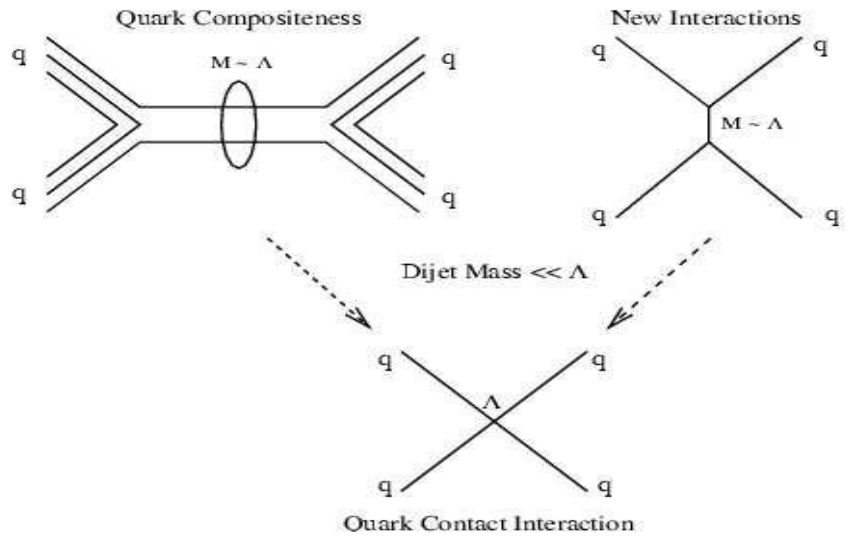


Figure 1.9: In the top diagrams, a s-channel illustrates the quark substructure (left) and t-channel represents exchange of a force carrier (right).The bottom Feynman diagram shows for contact interaction.

The experimental signature for contact interactions would be an increase in dijet production in the high-invariant mass region, but no resonant structure is predicted.

Chapter 2

LHC and the ATLAS experiment

2.1 Introduction

The Large Hadron Collider (LHC) at CERN, a particle physics research centre established in 1954, is the largest and highest energy particle accelerator in the world. It is designed to research on fundamental interactions, describe the behaviour of fundamental particles and new physics.

In this chapter, it is outlined the design of the LHC and ATLAS detector in section 2.2 and considered the 2.3 since I have been worked on the ATLAS. The parts of ATLAS, which are the inner Detector (subsection 2.3.1), the calorimetry (subsection 2.3.2), the muon spectrometer (subsection 2.3.3) are focused. Interesting events are selected using a trigger system (Section 2.4). The collected data is compared with theory prediction. To make theory-based predictions. Monte Carlo (MC) event generators (section 2.6) are used for running experiments and help preparation, for future experiments.

2.2 The LHC

In the collider, two high-energy proton beams are accelerated to 99.99999% of the speed of light in an almost circular tunnel spanning 27 kilometres. The

two beams travel in separate beam pipes, and collide with a peak crossing rate of 20 MHz, with bunches of protons interacting every 50 nanoseconds in 2012 at four different intersection points.

The superconducting magnet system steers the beam around the ring. Its dipole magnets are operated at 8.3 T which is more than 150,000 times the Earths magnetic field. In order to conduct electricity efficiently without resistance or loss of energy, the cables of the electromagnets' coils are made of an alloy of niobium and titanium, kept at a temperature -271.3°C. The LHC contains thousands of magnets, used for different purposes. While dipole magnets are used to bend the beams, the beams are focused by quadrupole magnets at regular intervals [36].

The LHC collisions were performed at an unprecedented centre-of-mass energy of 7 TeV for an instantaneous luminosity of $10^{34} cm^{-2} s^{-1}$ in 2010 and 2011. The proton energy in 2012 was increased to centre-of-mass energy of 8 TeV. The machine did shut down in 2013 and 2014, and then the energy of the accelerator increased to 13 TeV in 2015.

As illustrated in Figure 2.1, four main detectors are located in large caverns around the collision points of the LHC, A Large Toriodal ApparatuS (ATLAS), the Compact Muon Solenoid (CMS), LHCbeauty (LHCb) and Large heavy Ion Collider Experiment (ALICE). ATLAS and CMS have been built to investigate a wide range of physics and to independently confirm any new discoveries made [36].

2.3 The ATLAS detector

The ATLAS detector has been designed to understand matter, forces, the origin of mass and to discover new particles beyond the Standard Model such as supersymmetric particle or exotic dark matter candidates. ATLAS located at the LHC Point 1 is a complex detector with an inner tracker, a calorimeter system and external muon chambers. It is 46 m long, 25 meters in diameter,

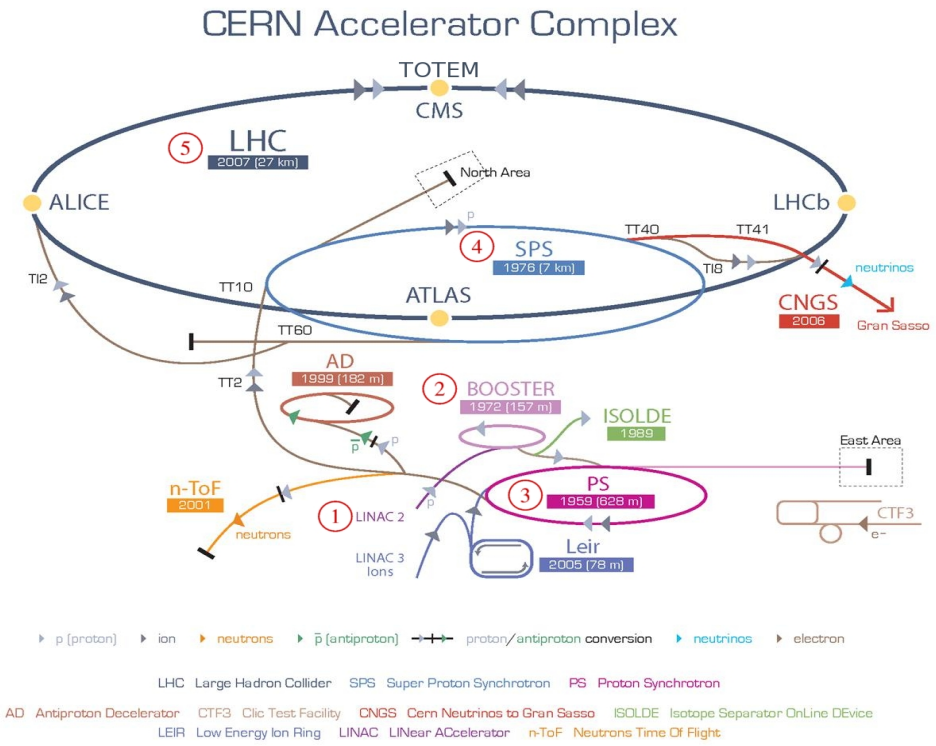


Figure 2.1: Overview schematic of the LHC represents the four main detectors and the two ring structure of the LHC [37].

and weights about 7000 tonnes; it has nearly 3000km of cabling.

In ATLAS, a right-handed coordinate system is used. The x-axis is aligned to the middle of the LHC ring, y-axis is aligned vertically upwards, and the z-axis points to the beam direction. The definition of pseudo-rapidity of particles is

$$\eta = -ln\left[tan\frac{\theta}{2}\right] \tag{2.1}$$

where the polar angle θ and azimuthal angle ϕ are measured with respect to the z axis. Pseudo-rapidity η is used instead of the rapidity, y, at higher energy because y tends to η in the massless particle limit. Rapidity depends

on particle E and momentum p_z in the z-axis as

$$y = \frac{1}{2} ln \left(\frac{E + p_z}{E - p_z} \right) \tag{2.2}$$

The magnetic field in ATLAS is comprised of a solenoid coil near the interaction point and a large toroidal magnet system in the Muon detector. ATLAS has a total of 11 large magnets: a central solenoid, eight external large magnets whose combination produces a toroidal field (hence the name), and two forward-backward toroids. These magnetic fields provide a measurement of the particles charge and their momenta.

There are three main calorimeter systems, ECAL (electromagnetic calorimeter), HCAL (hadronic calorimeter), and FCAL (forward calorimeters) as pictured in Figure 2.2. The general purpose of the calorimeters is to measure the energy of both charged and neutral particles, the showers from electrons and photons in the electromagnetic calorimeter and those from hadrons in the hadronic calorimeter. As muons pass through these layers of the detector as minimally ionising particles, the calorimeters are surrounded by muon chambers to measure the properties of these muons.

2.3.1 Inner Detector

In the inner detector charged particle are detected and, their momenta and charge are measured. It has a total length 7 m and a radius of 1.15 m, and is located inside the central solenoid magnet. It consists of silicon pixel detectors closest to the interaction point, the semiconductor tracker (SCT), and a straw tube transition radiation tracking detector (TRT) aligned with the beam direction as shown in Figure 2.3. The materials are relatively thin and low density because particle do not scatter too much and pass through. The detectors are comprised of the barrel and end-cap components for each of three sub-systems of the inner detector.

The silicon pixel detectors and the semiconductor tracker (SCT):

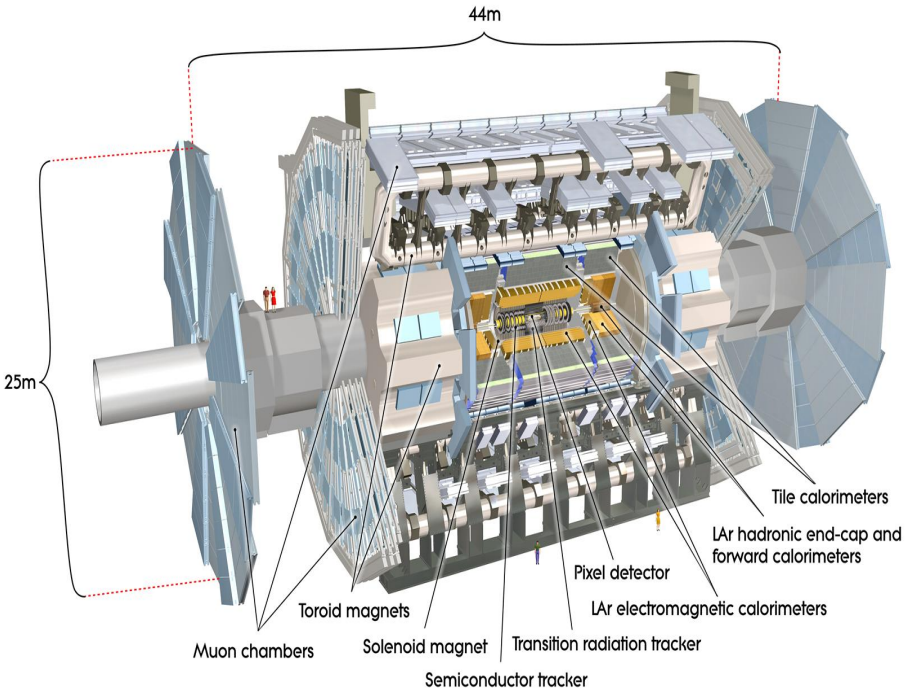


Figure 2.2: A cut-away view of the ATLAS detector [38].

The silicon pixel detector has a length of ∼1.3m and is made of 3 barrel layers near the beam pipe, 2x3 disks in the forward direction and readout channels. The semiconductor tracker consists of eight layers of silicon micro-strip detectors.Both detectors work in similar way and purpose. As charged particle goes through, it liberates e^-. This e^- moves to strip in SCT/pixel. It makes electric current as signal. The signal is used for information about track of particle [40].

The straw tube transition radiation tracking detectors: TRT is located between the solenoid and the silicon tracker. It consist of 'straw' drift tubes containing a Xe-CO_2-O_2 gas mixture. Charged particle pass through and it makes photon produce. Photon interacts with molecules in gas, free e^- moves to be measured as electric current. TRT also provides e^-/π^- separation.

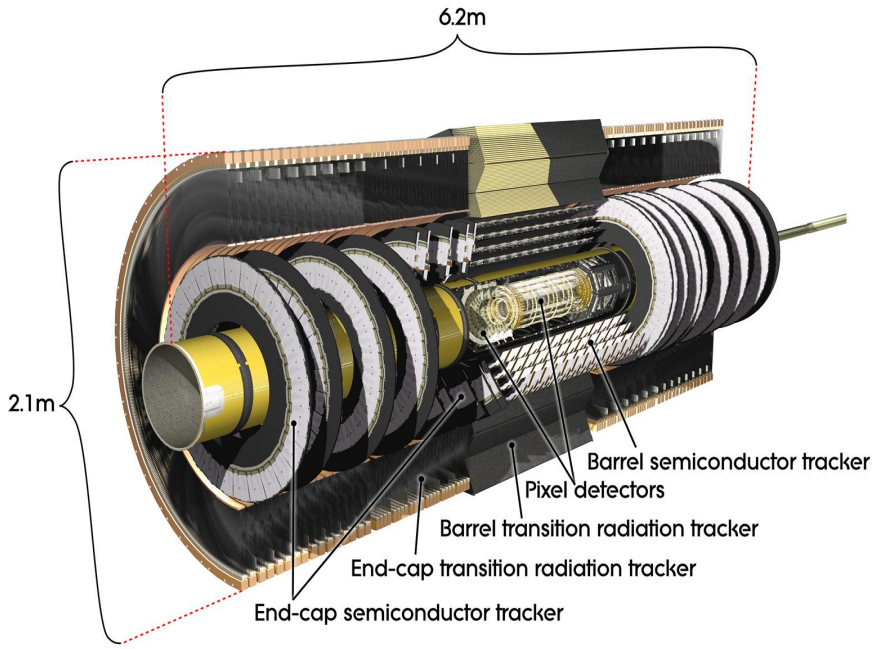


Figure 2.3: Overview of the ATLAS Inner Detector, with labels and dimensions [39].

2.3.2 The Calorimetry

The calorimeter system stops particles and, absorbs and measures energy of particles generated in a collision for particle identification. Neutral particles which cannot be detected in the tracking systems are only identified by measuring their energy deposits in the calorimeter.

The ATLAS calorimeters is sampling calorimeters and consist of absorber and active materials. When particle meet absorber, it interacts with material and start showers losing its energy. If absorber was not used, the calorimeter would be very large. The active material produces the detectable signal. There are three regions of the ATLAS calorimeter. The central region is defined within an η range of $|\eta| < 2.5$. The end-cap region is defined within the range $2.5 < |\eta| < 3.2$. Finally, the forward region is defined in the range $3.1 < |\eta| < 4.9$.

The calorimeter system measures both electromagnetic and hadronic show-

ers that start outside of the inner detector. Electromagnetic showers are much narrower than hadronic ones because less process are involved in their development, on the other hand EM showers penetrate much more material. Therefore, the electromagnetic calorimeter is placed inside of the hadronic calorimeter [41].

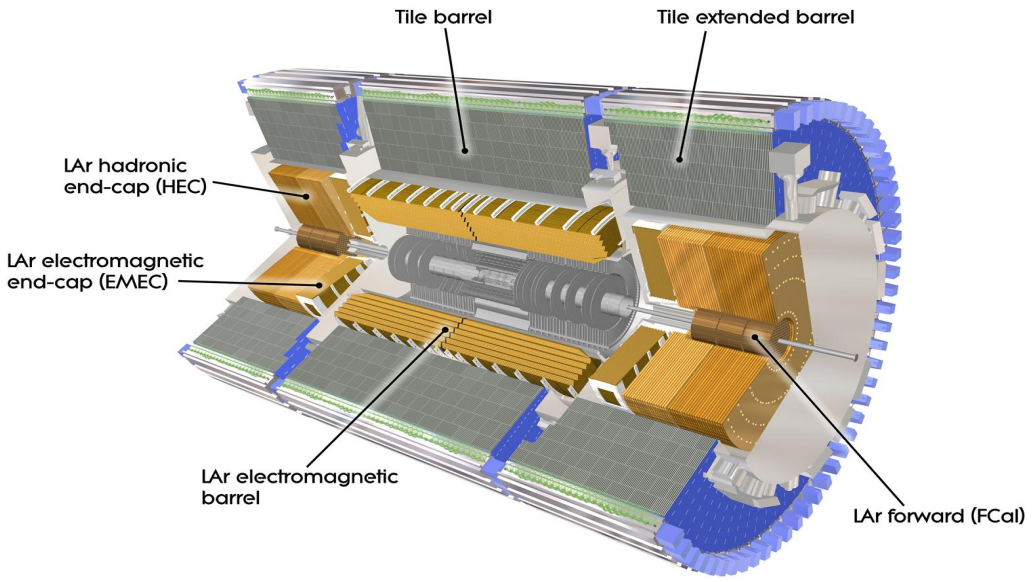


Figure 2.4: Cut-away view of the ATLAS calorimetry [41].

A layout of the calorimeter is shown in Figure 2.4. The thickness of the barrel of the EM calorimeter is more than 22 radiation length of the EM showers, X_0, (24 X_0 in the end-caps). The depth of the hadronic calorimeter is 9.7 nuclear interaction lengths, λ, in the barrel and 10λ in the end-caps. The amount of material as a function of pseudorapidity is shown in terms of radiation lengths in Figure 2.5.

Electromagnetic calorimeter: The Electromagnetic Calorimeter allows to measure the energy of particles such as electrons and photons. In addition to measuring the energy of electromagnetic showers, the ECAL provides discrimination between electrons and hadrons using the shower shape and pen-

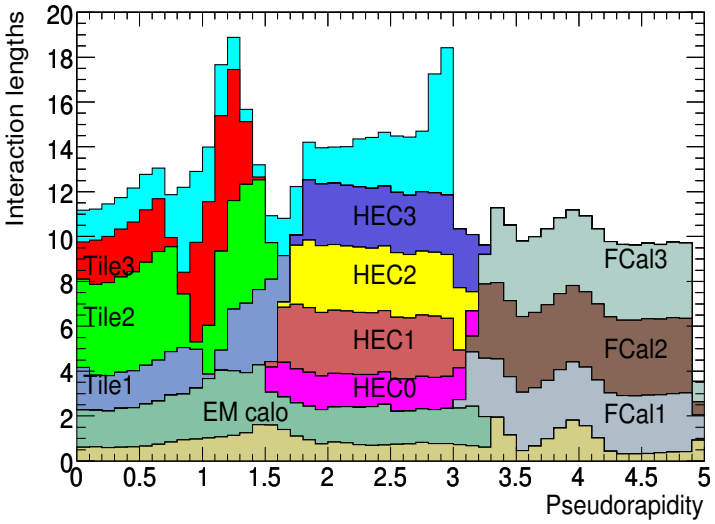


Figure 2.5: Cumulative amount of material in the ATLAS detector, in units of interaction length, λ, as a function of $|\eta|$. The coverage from each individual calorimeter component is shown separately, while the sections closest to the interaction point, which are not instrumented for calorimetry, are shown in brown [42].

etration.

The ECAL is a lead-Liquid Argon(LAr) detector with accordion-shaped electrodes and lead absorber plates over its full coverage. The reasons for the use of this technology are its radiation tolerance, intrinsic linear behaviour, and the stability of the response in time. The working principle of ECAL is that particles in the electromagnetic shower pass through material, it ionises atoms and negative charged e^- and positive charged ions. e^- drift through the LAr driven by electric field. The signal is derived from this e^- current. The current is digitised and recorded.

The ECAL consists of a barrel region around the inner detector cavity and two end-cap regions formed of outer wheels covering the region ($1.375 < |\eta| < 2.5$) , the inner wheels ($2.5 < |\eta| < 3.2$) and the forward region ($3.1 < |\eta| < 4.9$). The dead material of the central solenoid is in front of the ECAL, limiting its energy resolution.

Hadronic calorimeter: Hadrons such as protons, pions or neutrons are

detected in the hadronic calorimeter. There are four different regions in the hadronic calorimeter: the barrel ($|\eta| < 0.8$), the extended barrel ($0.8 < |\eta| < 1.7$), both of which are made of tile of scintillator(active material)/steel(absorber) sampling calorimeters, the hadronic end-cap (HEC; $1.5 < |\eta| < 3.2$), and the hadronic forward calorimeter, FCal, which uses LAr/W modules. The LAr(active material) hadronic calorimeters are made with a copper absorber rather than lead. The LAr technology is chosen because of its intrinsic radiation hardness at larger rapidities [42].

In hadronic calorimeter, particles interact with atom nuclei starting hadron shower in the absorber. This shower causes to radiate light and long optical fibres carry light to photomultiplier tubes to convert to electric currents.

2.3.3 Muon spectrometer

Muons penetrate through the calorimeters and fly towards the outermost part of ATLAS, the Muon detector. When muons pass through the spectrometer, they create electrical pulses in the wires. Muon positions can be found measuring the electrical pulse timing. The trajectories of muons are measured in order to identify their direction, their electric charge and their momentum. The magnetic field produced by the huge toroidal coils interleaved with the muon chambers allows the muon momentum to be measured more accurately than in the trackers [43].

It consists of thousands of long tubes filled with gas and thin read-out drivers, and is located in the outermost part of ATLAS. The spectrometer has one barrel and two end-cap components.

2.4 Trigger system

LHC is designed with a bunch crossing rate of ∼20 MHz, but the data recording system is limited to a maximal rate of about 1 kHz at the final High-Level Trigger, so most of the events produced in the collisions cannot be recorded,

and a very strong real-time selection is needed. Usually, low p_T events are considered to be less interesting, and are those more likely rejected. The system that decides which events to keep and which to reject is called trigger [44].

The trigger system consists of both hardware and software components. In ATLAS it is divided in three levels: Level-1 (hardware based), Level 2 and event filter (EF) which form the software-based High Level Trigger(HLT). A flow chart of the ATLAS trigger system is shown in Figure 2.6. Each of trigger levels have different selection criteria to allow events to pass to the next step at a rate that can be handled. If an event satisfies all criteria at any trigger level it is accepted. For some event topologies, even the requirement of passing all three trigger levels would lead to a too large rate. To maintain the rate to acceptable levels, these triggers are "prescaled", namely only a fraction of the events passing the trigger selections are actually stored for offline analysis.

A more detailed description of each trigger level follows:

Trigger Level-1: The Level-1 trigger uses information coming from the calorimeter and muon detectors. It needs to decide in ∼2 micro-seconds on whether events should pass or not to the next level. This time includes the latency of cables between the detector and the trigger logic. The output of Level-1 is a number of Regions of Interest (RoIs), containing muon-candidate tracks or clusters of energy depositions in the calorimeter, to be used by the HLT. The read-out drivers carry the information about the RoIs to the HLT; if there are no RoIs, then the event is rejected. Incoming data is held in memory through pipelined front-end electronics during the Level-1 decision time. Events containing high p_T leptons, jets, photons and high E_T^{miss} or E_T^{total} are identified through the L1 trigger, reducing the event rate from 20 MHz to ∼70 kHz.

High-Level Trigger: The High Level Trigger(HLT), Level-2 and EF, are software running on a CPU farm near the detector. The event rate is decreased to ∼6.5 kHz by the Level 2 using reconstructed electrons, jets, photons and muons. If a Level 2 trigger algorithm accepts an event, it passes to the EF.

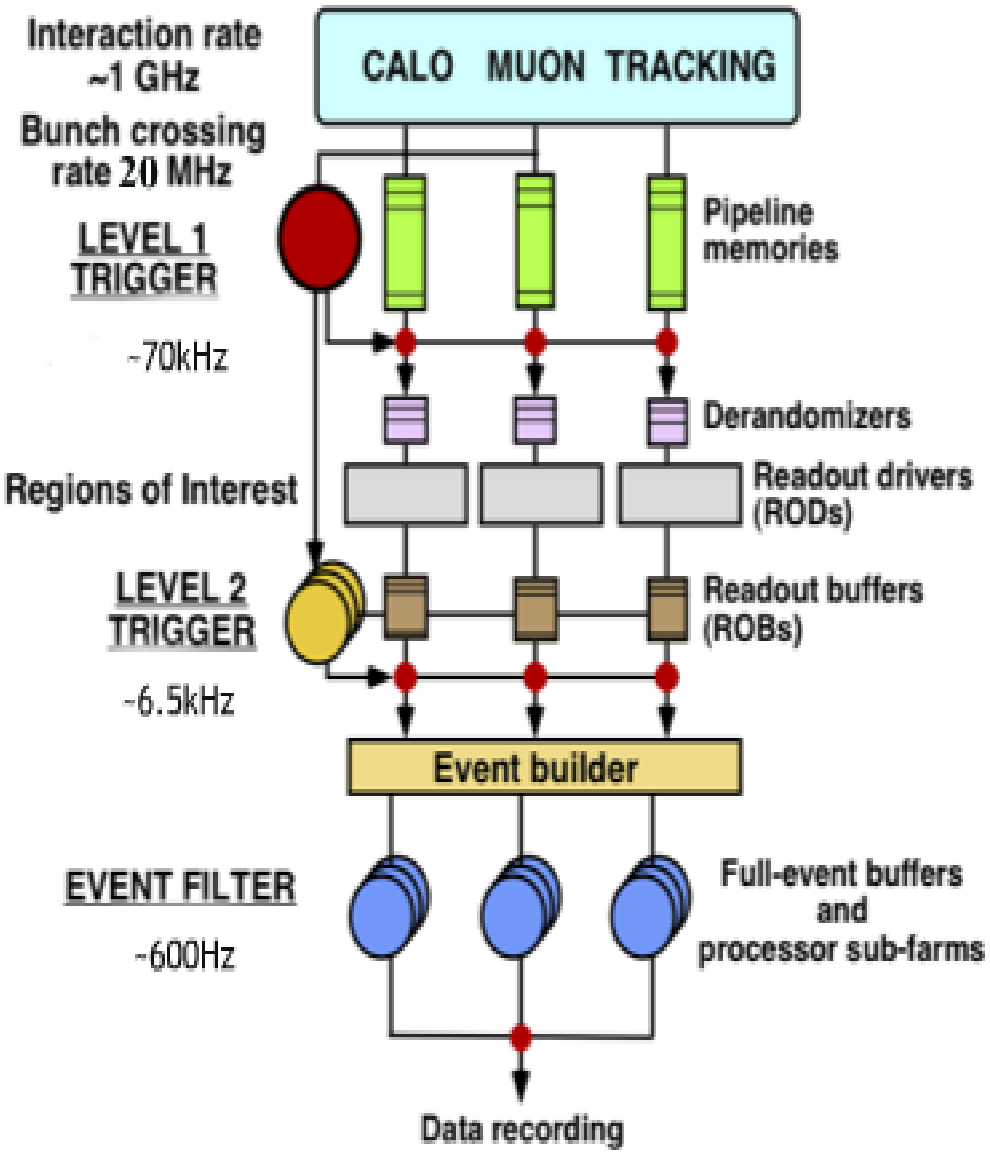


Figure 2.6: ATLAS trigger system [45].

This trigger reads out the full event information and makes a final decision about whether the event is stored or not on disk in ∼2 seconds. The events rate is reduced to ∼600 Hz.

2.5 Luminosity

The number of events occurring for a given process per unit of time is connected to the cross-section σ by the relation

$$\mathcal{L} = \frac{1}{\sigma}\frac{\mathrm{dN}}{\mathrm{d}t} \tag{2.3}$$

where $\mathcal{L}$ is called the luminosity. This quantity depends on the properties of the beam with the relation

$$\mathcal{L} = \frac{n_1 f n_2}{4\pi\sigma_x\sigma_y} \tag{2.4}$$

where f is the rotation frequency of the beam, n_1 and n_2 are the number of particles in each per brunch. σ_x and σ_y are respectively the RMS width of the beam in the horizontal and vertical directions [46]. To increase luminosity, the width of the beam can be reduced, or the number of protons/bunches in each bunch increased. Increasing luminosity results in more collisions, but if the number of colliding bunches is not increased, will result in increasing the number of collisions in the same bunch crossing (a phenomenon known as pileup), making the interpretation of the results more difficult. Figure 2.7 shows the luminosity delivered by the LHC and that recorded by the ATLAS detector during the 2010, 2011 and 2012 proton-proton data-taking periods.

2.6 Monte Carlo Generators

The Monte Carlo method gets its name from the use of random numbers, like the roulette games of the famous casino. To produce events directly comparable to data, two steps are needed: event generation and detector simulation [47]. In the generation step, particle level information is produced using theoretical and phenomenological models. The event generators provide the cross-section of various physics process in a defined kinematic region. Events are produced by randomly selecting from the available phase space using ran-

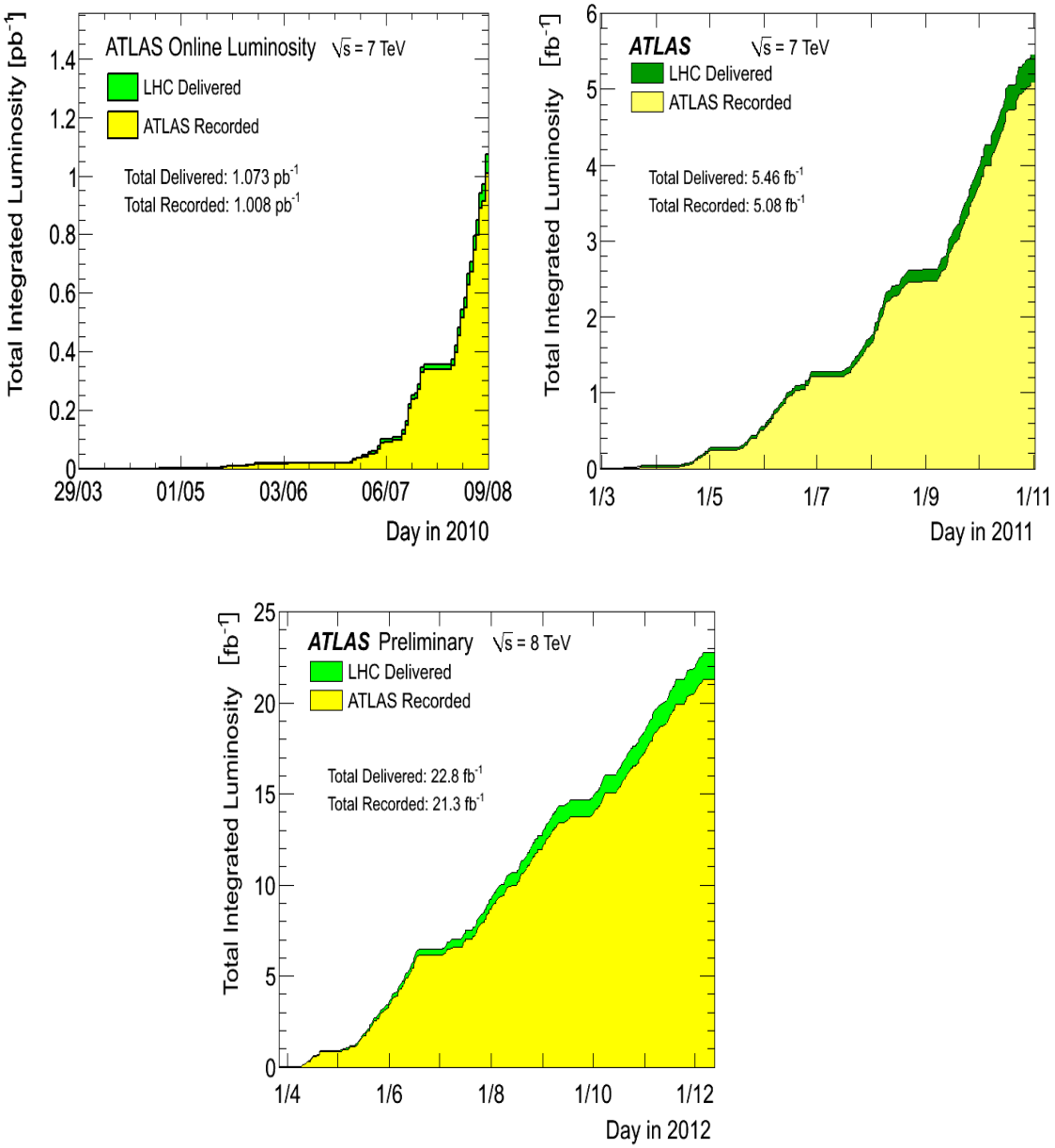


Figure 2.7: Cumulative luminosity versus time delivered to (green), and recorded by ATLAS (yellow) during stable beams and for pp collisions at 7 TeV centre-of-mass energy in 2010 (top left) and 2011 (top right) and pp collisions at 8 TeV centre-of-mass energy in 2012 (bottom) in ATLAS detector [46].

dom numbers.

The four main steps for simulating events using a generator are shown in Figure 2.8; the hard scatter, initial and final state radiation, hadronisation and the subsequent hadronic decays.

MC event production starts with a basic subprocess coming from a highly energetic collision of initial-state particles. For instance, a pair of gluons or a light quark-antiquark from incoming protons collide and generate a top quark-antiquark pair. The generator evolves particles until a stable final state, with a proper lifetime longer than $ct > 10$mm. The particles at hadron level are propagated through a detector simulation program based on the GEANT4 [48] code within the ATLAS simulation framework. This software simulates a model of the ATLAS detector, including effects such as particle ionisation in the trackers, energy deposition in the calorimeters, intermediate decays and interactions with non-detecting material.

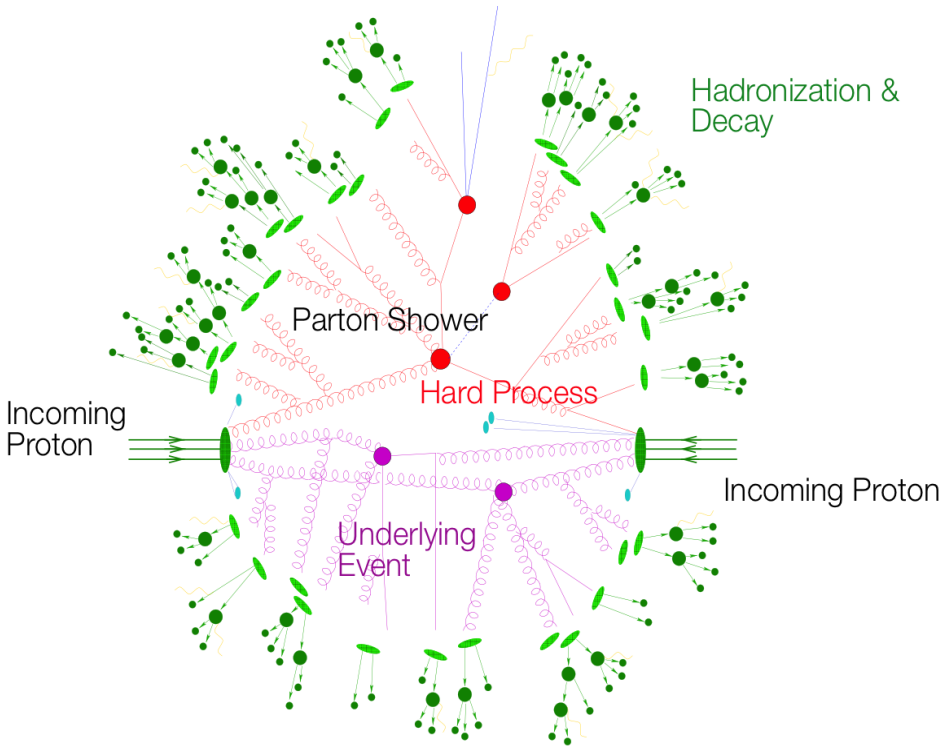


Figure 2.8: Sketch of a hadron-hadron collision as simulated by a Monte-Carlo event generator [47].

Several MC generators exist for the vast majority of processes of interest in the LHC, differing in the order of perturbative expansion, in the approach to multiple final states, and the approximation made to account for the missing calculations. Also different can be hadronisation, the treatment of Multiple Parton Interactions which is semi hard interaction comparing to first one in same collision, the colour reconnection models and parton showering. The generators can be interfaced with different sets of PDF.

2.6.1 Pythia

Pythia is a general-purpose event generator and produces $2 \Rightarrow 2$ LO(Leading Order) pQCD matrix elements followed by a p_T-ordered parton shower based on the string model. Currently, two versions of Pythia are used: Pythia 6 [27], the Fortran version, and Pythia 8 [49], the C++ implementation. They differ among other things in their approach to parton shower: Pythia 6 can use with virtuality which is set by some momentum transfer scale Q of the hard subprocess or transverse momentum as the evolution variable, while Pythia 8 can generate the dipole showering which also takes into account soft gluon coherence. The Lund string model is applied in both of them to describe hadronization and the underlying event which occurs mainly from multiple parton interactions. A set of tuning parameters can be modified to reproduce measurements of jets shapes, azimuthal decorrelations etc. performed at previous experiments like CDF and D0.

2.6.2 Alpgen

Alpgen is a generator for producing LO matrix elements with hard multi-parton process up to six partons in the final state. It is used for final states where several well separated hard jets can be present. Events are produced at the parton level, taking into account the full information on their colour and flavour structure, allowing for the evolution of the partons into fully hadronised

final states. These partons have to be interfaced with a subsequent parton-shower and hadronisation code, like Pythia [50].

2.6.3 Next-to-Leading-Order Theoretical Predictions

Theorists are able to calculate almost all processes at NLO in powers of α_s. In this thesis, NLOJet++ is used for measurement of the inclusive-jet cross-section using a low μ Run.

NLOJet++

The NLOJET++ program is not a proper MC event generator, but an analytical code that can explicitly calculate leading and next-to-leading order cross-sections, containing 2→2 virtual-loop correction and 2→3 real emission. It can also generate events, but only at parton level, requiring the use of non-perturbative corrections to compare these results with data.

This code can be interfaced with the APPLgrid software [51], which stores the perturbative coefficients of the theoretical calculation in a two-dimensional grids, allowing fast calculations of the NLO QCD cross-section for any parton density function (PDF), and any value of the strong coupling (α_s), and any value of the renormalisation and factorisation scale. It therefore allows the computation of the three main sources of uncertainty for a theory prediction.

Several PDF sets such as CT10 [52], HERAPDF15 NLO_EIG, MSTW2008nlo68cl [53], MMHT2014 PDF [54], and NNPDF21_100 [55] are used in the measurement of the inclusive-jet cross-section. The uncertainties on the strange and bottom quark masses is included in MSTW 2008, while NNPDF21 and HERAPDF15 also have the uncertainties of strange-quark fraction and a Q^2 cut.

Non-perturbative corrections

As mentioned above, NLOJet++ calculations can only provide parton-level cross sections. To compare these prediction with data, the output from NLOJet++ should be corrected for hadronization and underlying events effects,

using the so-called non-perturbative corrections. They are derived from LO MC generators like Pythia, and are applied on a bin-by-bin basis. The correction is done according to the following steps:

- a leading-order Monte Carlo such as Pythia or Herwig produces final distributions with and without hadronisation and underlying event
- the ratio between final distributions with and without hadronisation and underlying event is taken
- The ratio between these two distributions is then multiplied by the Next-to-Leading Order distribution from NLOJet++.

In order to allow an estimation of the uncertainty, several different leading order Monte Carlos codes and sets of hadronisation parameters are used in the determination of these non-perturbative corrections. The main Monte Carlo tunes used in this analysis are pythia6_P2012, pythia6_AUET2BLO, herwigpp_CTEQEE4 [56].

Chapter 3

Measurement of the inclusive-jet cross-section using a low μ Run

3.1 Introduction

Jet production is the dominant high transverse-momentum process at the LHC. Measuring the inclusive cross-section is important to study high order QCD , to extract information about the structure of the proton, the strong coupling constant α_s and physics beyond the Standard Model. With the increased energy and integrated luminosity of the 2012 run, the LHC allowed testing previously unexplored regions of phase space for jets with high p_T. However higher luminosity came at the price of an increased number of collisions per bunch-crossing, and in 2012 for each high-p_T event there were on average about 30 simultaneous other proton collisions, making jet calibration and the determination of a proper jet energy scale (JES) more difficult, especially at low transverse momenta.

Thanks to the increased accuracy of jet calibration in energy scale, the jet cross section measurement in 2012 is expected significantly more precise with respect to the previous measurement at 7 TeV. However, due to presence of large pile-up uncertainties at low p_T jet, it is difficult to improve the precision below transverse momenta of about 100 GeV. For this reason, normal high luminosity run have been used in 2012 to measure inclusive jet cross-section

while the low-p_T region, with jet transverse momenta down to 25 GeV, has been measured using the special low μ [1] runs when pileup was much smaller. This thesis work describes the low p_T analysis performed on the special runs, that has been combined with the high transverse momentum part measured by Gagik Vardanyan.

3.2 Cross-section Definition

Jets produced in proton-proton collisions at $\sqrt{s} = 8$ TeV are reconstructed using the anti-k_t clustering algorithm, with both jet size parameters of 0.4 and 0.6. The recombination scheme, which is a set of a rules for obtaining the four-momentum of a jet from its constituents, is taken. In the measurements, experimental detector effects are removed using Monte Carlo simulation. The measurements are defined by the inclusive jet cross-sections obtained from the stable particles entering the ATLAS detector. The stable particles are defined as those with a proper lifetime longer than $ct > 10$mm after the hadronisation process. This definition also includes muon and neutrinos. All jets within an acceptance of $|y^2| < 3.0$ and p_T >15 GeV are considered. The double-differential cross sections for inclusive jets measured as a functions of jet p_T and rapidity is defined as:

$$\frac{d^2\sigma}{dp_T dy} = \frac{1}{\epsilon L}\frac{N}{\Delta p_T \Delta y} \tag{3.1}$$

where L is the integrated luminosity, N is the number of events in the jet p_T and rapidity bin, and ϵ is the selection efficiency. The cross-section is measured in six rapidity bins of $|y|$ between 0.0 and 3.0, each with a width of 0.5. The measurement are unfolded to particle level (consisting of final-state particles before detector effects), accounting for the effects of hadronization and the underlying event in addition to the hard scatter.

[1] μ is the average number of collisions per brunch crossing.

[2] The rapidity is $y = \frac{1}{2} ln\left(\frac{E+p_z}{E-p_z}\right)$ where pz is the momentum of the jet in the direction of the beam line and E is the energy of the jet.

3.3 Data-set used

The jet cross section measurements uses the 2012 data-set of proton-proton collisions with data periods between A2 and A6. Two different runs have been used in the analysis. They will be called the no pile-up run, with $\mu = \sim 0.5$ and the low μ run, where $1.4 < \mu < 1.6$ and $4 < \mu < 6$. The accelerator was run on the no pile-up run at reduced instantaneous luminosities, in correspondence to a Van der Meer scan [57], where the beam has been swept transversally to measure its profile, and eventually calibrate the luminosity; this run was taken on April 12, 2012. A low μ run was taken on 16-17 of April 2012. The total luminosity of the no pile-up and low mu runs is 193 nb^{-1}. The preliminary uncertainty on this luminosity is $\pm 10\%$ based educated guesses for the no pile-up run, while it is $\pm 2.8\%$ for the low μ run [46]. The data using in this analysis are required to have good detector status and data quality defined by the standard Good Run List (GRL)[3] for period A2 to A6 [58].

The run 200805 is called no pile up run is so small ($\mu = \sim 0.5$) that no additional pile-up interactions are present. For that special run, the calorimeter noise thresholds used to build the topological cluster have been decreased and correspond only to the electronic noise. A special Monte Carlo simulation for this run has been produced at same conditions with data.

The runs 201351 and 201383 are called low μ since some pile-up is present but much smaller than during the rest of the year. The nominal calorimeter conditions where the noise thresholds correspond to the electronic noise are added in quadrature to the pile-up noise. The inclusive jet cross-section is measured in the transverse momentum range where 25-45GeV using the no pile-up run, while the p_T range between 45 and 100 GeV is measured using the low μ run. The reason of using no pile-up run between 25-45GeV is lack of statistic at $p_T > 45$GeV. Therefore, the low μ run has been added $p_T > 45$GeV to have sufficient statistics.

[3]Good Run List files tells which luminosity blocks to exclude to have "good" data.

3.4 Event Selection

The requirements for the events to be considered for the inclusive jet cross-section measurement are described as follows:

Primary vertex :

Events from cosmic-ray muons and other non-collision backgrounds are removed by requiring that good events must have at least one primary vertex consistent with the beam-spot position, and that at least two tracks are associated to it. This requirement is necessary for identifying pile-up vertices for the jet calibration [59].

LAr and Tile read-out error requirements : Events are rejected if they contain errors in the read-out of the Lar and Tile calorimeters or noise bursts in the LAr calorimeter, as recommended by the ATLAS data preparation group [60].

Incomplete data events: Events are rejected using the coreFlags[4] variable, if they have incomplete or corrupted data.

Trigger Selection: Events must satisfy trigger selection. The integrated luminosity per used trigger is summarised in Table 3.1.

Trigger	Luminosity μb^{-1}
EF_mbMbts_2_NoAlg[5]	113.54
EF_j15_a4tchad	54314.8
EF_j45_a4tchad	93252.8

Table 3.1: Trigger selection, Luminosity for anti-k_t jets with R = 0.4 and R = 0.6 are shown

[4]The coreFlags is defined as event flag for each core sub-detector.

[5]**EF_mbMbts_2_NoAlg:** A minimum Bias trigger provides the primary requirement for selecting events from real collisions with the smallest bias. It requires at least two hits in the MBTS (Minimum Bias Trigger Scintillators) detectors, placed before the calorimeter endcap.

EF_j15_a4tchad: requires at least a jet with transverse energy (E_T) above 15 GeV

GRLs cut: The run should belong to a GRLs to have good data.

The number of events remaining after the above selection criteria is shown in table 3.2 for the no pile-up run and 3.3 for the low-mu run. The larError, tileError, coreFlags cuts are negligible in the both runs.

No Pile-up Run	
Event-level cut	Number of Event
Total	1.161687e+07
larError, tileError,coreFlags	1.161687e+07
primary vertex requirement	9.42992e+06
Good Run List (GRL)	1.14828e+07
Trigger	9.0433e+06
After all cut	8.18242e+06

Table 3.2: Events for anti-k_t jets with radius parameter R = 0.4 are applied selection criteria for no pile-up run.

Low Pile-up Run	
Event-level cut	Number of Event
Total	9.540106e+06
larError, tileError,coreFlags	9.540106e+06
primary vertex requirement	9.52437e+06
Good Run List (GRL)	9.5372e+06
Trigger	7.73205e+06
After all cut	7.72355e+06

Table 3.3: Events for anti-k_t jets with radius parameter R = 0.4 are applied selection criteria for low μ-run.

3.5 Jet Selection

The energy measurement is not accurate in some regions such as the transition between the barrel and end-cap, and the other problematic calorimeter regions. Jets in problematic calorimeter regions are known as "ugly", and not well

at the EF. "a4tchad" means that jets are reconstructed using the anti-k_t algorithm with topological cell energy clusters as signal input with a radius of R=0.4

EF_j45_a4tchad: requires a jet the same characteristics as the previous case, but transverse energy (E_T) above 45 GeV.

measured. The energy coming either from TileGap3, the transition region between the barrel and end-cap, or from known dead cells, which are assigned an energy value based on the values of their neighbouring cells cause to exist "ugly" jets.

Moreover, jets called as "bad" are not associated to real energy deposits in the calorimeters. Some hardware problems (HEC spike, EM coherent noise) LHC beam conditions and cosmic-ray showers can cause bad jets [61]. In this measurement, the "BadLoose" and "Ugly" cut are used to remove these kind of jets for avoiding the above problems and missing tails due to detector failures. The inefficiency is a very small(<0.6%) and can be neglected [62].

Corrupted Tile Data Veto: The Tile calorimeter suffered from frequent power supply trips during data taking. A dedicated procedure to identify such events is provided by the Tile calorimeter system and these effects are corrected during the offline reconstruction. However, for a short period of time (run 211620-214553) and for one particular channel (A2 of LBC24 at $\eta = -0.15$, $\phi = 2.3$) this algorithm did not work properly. The problematic region is vetoed for the higher mu run on the jet level. This specific correction is not applied for the low μ run and the no pile-up run because the effect of the remaining module on jet with low p_T on these two runs is negligible [63] and there is no no module mask as seen in Figure 3.1.

Jets in both data and MC have to pass these two requirements for both the no pileup and low-mu data-set. Additionally, p_T and $|y|$ cuts are applied as in Table 3.4 for the no pile-up run and 3.5 for the low μ run.

3.6 Jet Reconstruction and Calibration

Jets are reconstructed using the anti-k_t algorithm with R = 0.4 or 0.6 (where R is the radius of the circle in y-ϕ phase space) using the FASTJET soft-

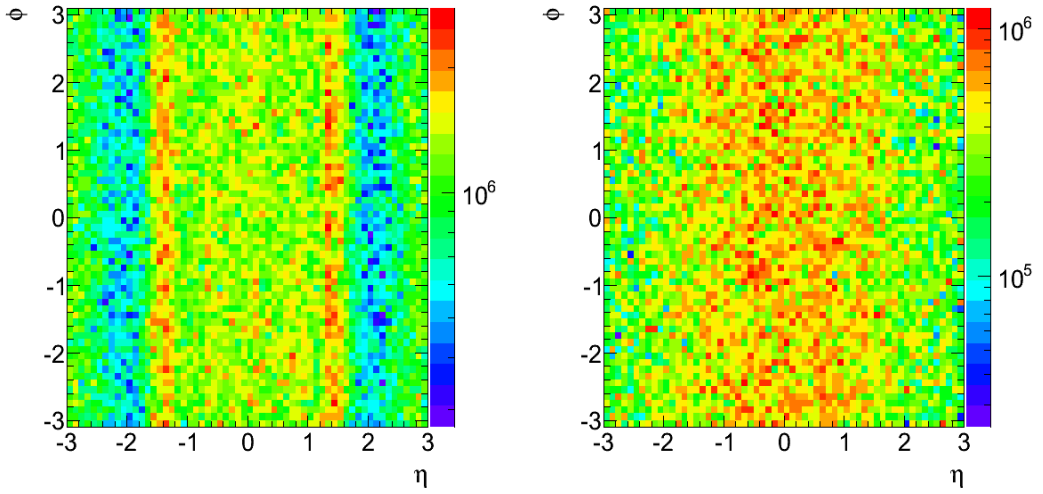


Figure 3.1: The jet pseudo-rapidity and the jet azimuthal angle distribution using reconstructed jets after event and jet selection cuts in data in R=0.4 for no pile-up run (left) and low μ run (right).

No Pile-up Run	
Jet-level cut(after event requirements)	Number of Jets
Total jets passing the cuts	2.00796e+07
No badloose/ugly jets	1.99599e+07
Jets in $\|y\| < 3.0$	1.45608e+07
Jets with $p_T > 15$ GeV	391415
Jets after cuts	312582

Table 3.4: Jets for anti-k_t jets with radius parameter R = 0.4, are applied jet selection criteria after event requirements for no pile-up run.

ware for both data and MC simulation. In ATLAS, jets are first reconstructed from three-dimensional topological clusters at the electromagnetic (EM) scale which corrects the jet energy to the energy deposited by an electromagnetic shower. Each topo-cluster is built from a seed calorimeter cell with $|E_{\text{cell}}| > 4\sigma$, where σ is the RMS of the noise of that cell. Adjacent cells are iteratively added to the topo-cluster if they satisfy the condition $|E_{\text{cell}}| > 2\sigma$. As a last step, the topo-cluster defined from the outer layer of all surrounding cells and added. To improve the energy resolution for each topo-cluster, a local calibration(LC) based on the topology of the calorimeter energy deposits is used. Extra corrections are implemented reconstructing the energy deposits of elec-

low μ Run	
Jet-level cut (after event requirements)	Number of Jets
Total jets passing the cuts	9.44021e+07
No badloose/ugly jets	9.4000016e+07
Jets in $\|y\| < 3.0$	8.17296e+07
Jets with $p_T > 15$ GeV	1.48595e+07
Jets after cuts	1.29598e+07

Table 3.5: Jets for anti-k_t jets with radius parameter R = 0.4, are applied selection criteria after event requirements for low μ run.

tromagnetic calorimeter showers. Additional corrections are applied for dead material, out-of-cluster losses for pions, and calorimeter response for hadronic showers [62, 64]. For purpose of jets reconstruction, calorimeter clusters are considered as massless particles with energy $E = \sum E_{\text{cell}}$ and position at the energy-weighted barycenter of the cells in the cluster and originating from the geometrical centre of the ATLAS detector.

Jets are then calibrated to account for detector effects (i.e energy loss in dead material in front of the calorimeter or between calorimeter segments) and to compensate for the lower calorimeter response to hadrons than to electrons or photons using constants derived from MC, and depending on jet transverse momentum and rapidity. The four-momenta of the reconstructed jets are calibrated using the actual vertex position and additional number of vertexes to get rid of pile-up effects. The following description gives detail about the corrections applied by the standard calibration tool [65].

- **Pile-up energy correction:** An offset is applied to correct jet energy from the additional contribution due to pile-up events. There are two steps to subtract additional energy because of pile-up interactions within the same bunch crossing (in-time)[6] due to events occurring in the same

[6]**In-time:** Additional proton-proton collisions in the same bunch-crossing as the collision of interest.

Out-of-time pile-up: additional proton-proton collisions in bunch-crossings just before and after the collision of interest. When detectors are sensitive to several bunch crossings or their electronics integrate over more than 25 ns, these collisions can affect the signal in

bunches and in near-by bunch crossings (out-of-time) due to events occurring in previous brunches [59].

  - **Event-by-event pile-up energy correction:** At first step, the correction is derived from MC. It is implemented as a function of the jet p_T density (ρ) per unit area and area of anti-k_t jet (A)[7] is found for each event [66]. This correction is not used for the no pileup run due to the non-existence of pile-up.
  - **Residual pile-up energy correction:** After applying the jet area correction, the second term removes the residual average p_T induced by pile-up interactions. This off-set correction depends on the expected number of additional collisions in the same crossing, μ, and the total number of reconstructed primary vertices in the event NPV(Number of Primary Vertex).

- **Jet Direction Correction:** The jet direction is changed to account for the fact that the actual collision does not exactly occur in the geometrical centre of the detector.

- **Jet energy correction:** To avoid instrumental effects (calorimeter non-compensation, additional dead material, out-of-cone effects), a correction is applied to the energy and the position of the jet. Additionally, the jet energy is scaled to match that of particle-level in the MC simulation.

- **Jet structure-based corrections:** A "Global Sequential calibration" (GSC) [62] correction is applied for fluctuations in the jet particle content using the topology of the calorimeter energy deposits and of the tracks associated to jets. It is found that GSC does not change the jet

the collision of interest.

[7]The jet area (A) is a measurement of the susceptibility of the jet to pile-up with jet-by-jet in the event

energy in this measurement, but improves the resolution and reduces the uncertainties. Its other impact is the sensitivity to jet fragmentation effects such that differences between quark- or gluon-initiated jets is reduced.

- **In-situ corrections for differences between data and Monte Carlo simulation:** All previous corrections are based on MC, and they are cross-checked using in-situ techniques to also correct for residual differences with data. In situ corrections are derived by comparing the results of Z+jet,γ+jet, dijet and multi-jet balance techniques. The p_T balance between a photon or a Z boson and a jet is used to get the agreement of the jet scale in the data with the one in the Monte Carlo [67].

3.7 Trigger Strategy

Data events are collected using different trigger requirements, where jets are selected according to their p_T and pseudorapidity η. The central jet triggers, which cover the pseudorapidity region $|\eta| < 3.2$ are used in this measurement. The offline cut requirement is that jets have $|y| < 3.0$. Therefore the jet energy is fully measured by the central trigger towers. The trigger selection is not used for the Monte Carlo simulation.

There are several independent trigger chains with different transverse energy (E_{T}) thresholds defined by the presence of a jet with sufficient E_{T} at the electromagnetic (EM) scale, i.e. the sum of the energy in the two calorimeters, without corrections. To reduce the enormous rates of jet production, some of the jet triggers are prescaled, meaning that only a (usually small) fraction of the events passing the trigger requirements are actually stored. For instance, EF_j15_a4tchad is a prescaled trigger.

To avoid having to correct rapidly-rising efficiencies, triggers are used in the region where their efficiency has a flat dependence on the jet transverse momentum, and is higher than 99.5%. For a given trigger, the trigger with the im-

mediately lower threshold is used to measure the trigger efficiency curve (bootstrap method). The efficiency is defined as the ratio between the number of jets passing a given p_T threshold divided by the number of jets triggered by the fully efficient monitor trigger. To calculate the efficiency of EF_j45_a4tchad, EF_j15_a4tchad is used as the fully efficient monitor trigger at $p_T >\sim 25$GeV.

In order to calculate the efficiency of the lowest threshold (EF_j15 a4tchad) trigger, the minbias trigger has been used as a reference. This trigger is 100 % efficient in the jet transverse momentum region of the EF_j15_a4tchad turn-on curve, so it will be an unbiased reference. Even if the minbias trigger is present in the menu used by the low-mu run, there are no events passing it. It is assumed that the trigger turn-on curves for the no-pileup and low-pileup runs are the same, and the efficiency is computed on the latter. The efficiency of the EF_j15 trigger has been computed requiring jets to pass the conditions $p_T > 15$ GeV and $|\eta| < 3.2$ where the quantities computed by the EF system are used in the cut. If at least one jet in the event passes these requirements, the event is triggered. At second step, the p_T distribution for jets passing this trigger is defined as the numerator, and the p_T distribution for all jets as the denominator. Figure 3.2 shows the trigger efficiencies for the EF_j15_a4chad and EF_j45_a4chad triggers in the central and forward region, for anti-k_t jets with $R = 0.4$. A fit is applied to the trigger efficiency curve using the function:

$$f(x) = a\ \mathrm{ERF}(p_\mathrm{T} - b)/c),$$

where ERF is the standard error function[8] and a, b, c are free parameters that are connected respectively defined to the maximum efficiency in the plateau region, the p_T value where the efficiency reaches half its maximum value, and the slope of the turn-on curve.

The trigger plateau is consistent with 100% and no correction is applied to

[8]The error function is defined: $(2/\sqrt{\pi}) \int \exp(-t^2)\ dt$.

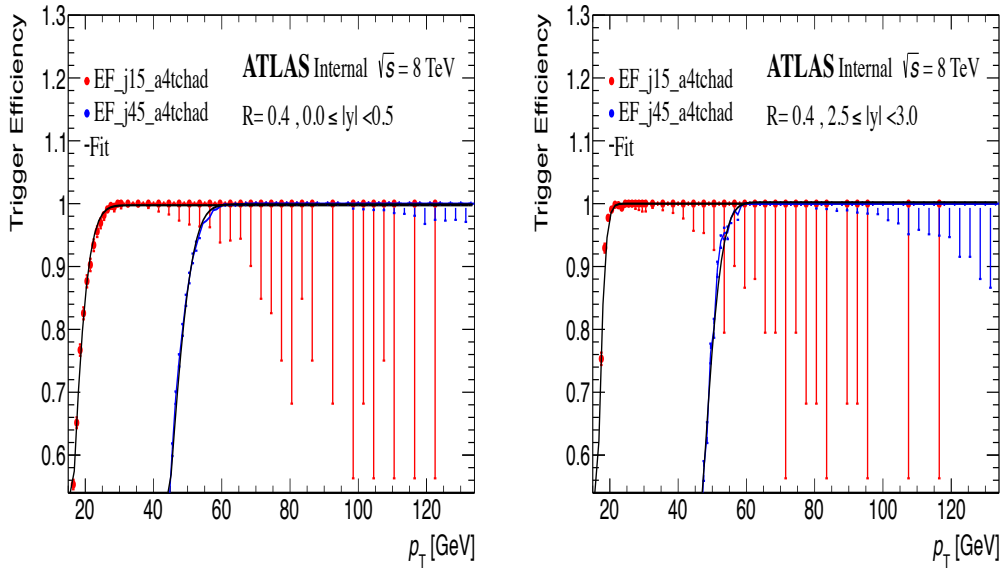


Figure 3.2: The efficiency of EF_j15_a4chad and EF_j45_a4chad trigger using anti-k_t jets with R=0.4 at $0.0 \leq |y| < 0.5$ and $2.5 \leq |y| < 3.0$

the efficiency in the regions of interest for the cross-section measurement. Each trigger is used in a p_T region where is fully efficient. Each event is weighted by the average effective luminosity recorded by each trigger. Table 3.6 summarises the triggers used for each of the jet p_T bins of the cross-section measurement.

Run Number	Trigger	p_T range (GeV)
200805	EF_mbMbts_2_NoAlg [68]	15-45
201351	EF_j15_a4tchad	45-85
201383	EF_j45_a4tchad	> 85

Table 3.6: The trigger chains used for each run in the no pile-up and low μ data-set for jets with $R = 0.4$. Indicated is also the transverse momentum range of interest for that trigger.

The triggers used in the high-μ runs are shown in Table 3.7

Trigger	Luminosity (pb^{-1})	p_T ranges (GeV)
EF_j360_a4tchad	20277.0	>478
EF_j280_a4tchad	1164.7	376-478
EF_j220_a4tchad	261.4	290-376
EF_j180_a4tchad	78.8	240-290
EF_j145_a4tchad	36.3	216-240
EF_j110_a4tchad	9.8	172-216
EF_j80_a4tchad	2.32	134-172
EF_j55_a4tchad	0.44	100-134

Table 3.7: The trigger chains used for each run in the high mu data-set for jets with $R = 0.4$. Indicated is also the total integrated luminosity and the p_T range where the trigger is applied.

3.8 Detector level Comparison

The unfolding of the experimental data, described in the next section, has been performed using the Pythia8 Monte Carlo based on the AU2 tune; this Monte Carlo has been interfaced with the CTEQ6L1 pdf set for the no pile-up run and with the CT10 pdf set for the low μ run.

This choice of Monte Carlo corresponds to the standard ATLAS conditions for 2012. Calorimeter cell noise thresholds corresponding to high μ run are used for the low μ run. For the no-pileup run, a dedicated Monte Carlo simulation has been used, with calorimeter noise thresholds set to near zero, corresponding on the conditions used in data. The comparison between data and MC are shown at detector level in Figure 3.3 for both separately no pile-up and low μ runs. As expected from a leading order Monte Carlo, simulation does not describe the absolute normalisation. However, the shape of the distribution is reasonably well reproduced, apart from the high-p_T region, not used in this analysis. The data to Monte Carlo ratio is relatively smooth as a function of the jet pT. It can be noticed that while measurement for the no-pileup run has been extended down to a transverse momentum of 25 GeV, its statistical uncertainties become large in the 100 GeV region.

Figure 3.4 compares the transverse momentum distributions from the Monte Carlo simulations of the zero-pileup and the low μ runs at detector level. Two

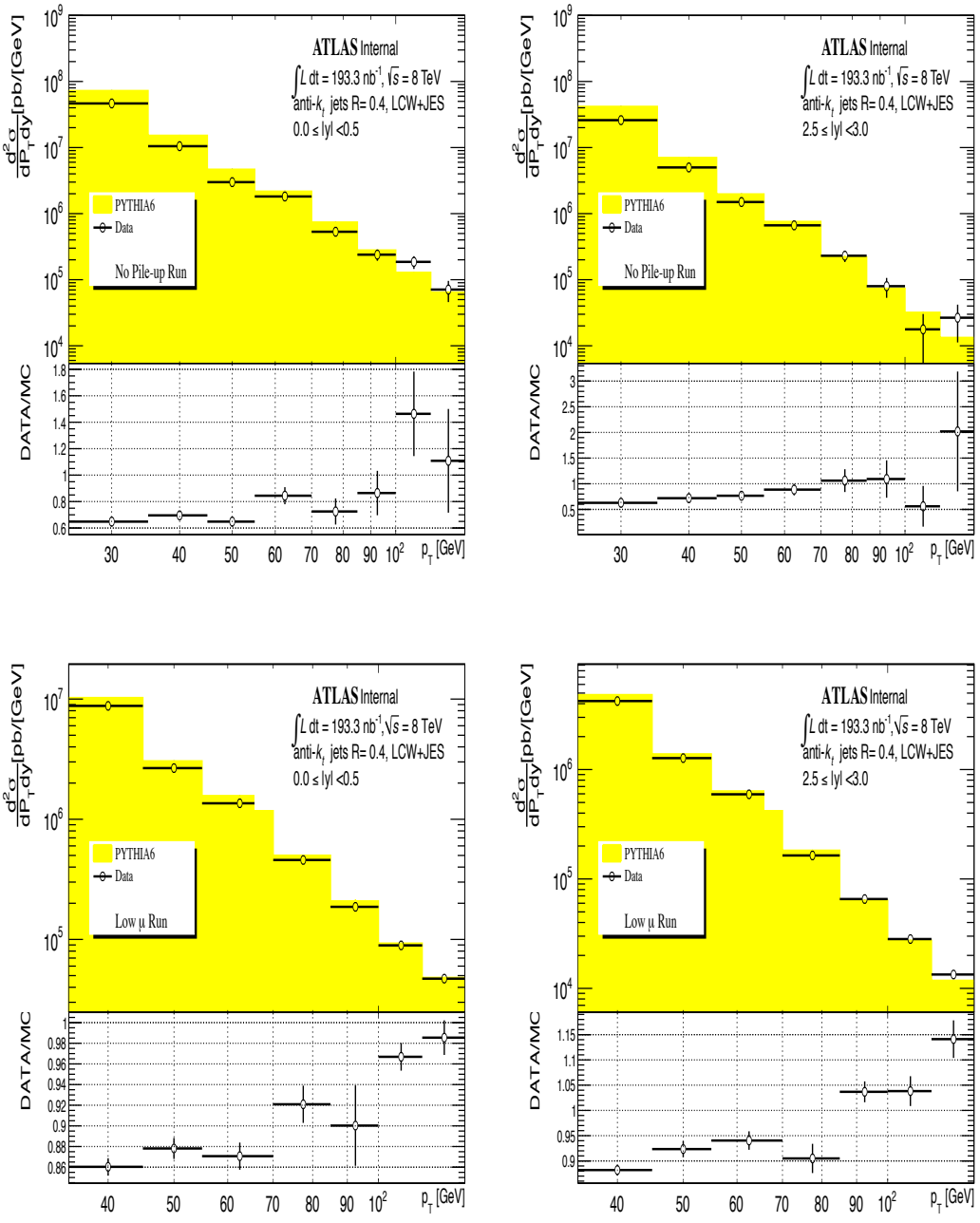


Figure 3.3: Comparison of reconstructed jet p_T distributions in data to the ones in the baseline MC simulation for anti-k_t jets with R = 0.4 in the regions $0.0 \leq |y| < 0.5$ (left) and $2.5 \leq |y| < 3.0$ (right), for no pile-up (top row) and low μ run (bottom row). The number of jets is normalised to the total integrated luminosity. In the lower part of the figure the ratio between data and MC simulation is shown.

samples, produced at different momentum transfer, are combined: JZ1W[9] and JZ2W; their merging point occurs for transverse momenta around 70-85GeV. The comparison between the jet transverse momentum distributions of the two runs in data is shown in the same figure. The ratio plot at the bottom indicates discrepancies of the order ∼10% in the central η region, and of about 20% in the forward region. These discrepancies are similar to what is observed comparing the Monte Carlo datasets simulated with different noise thresholds, so some of them are expected to cancel out after the unfolding procedure has been applied. It is also worth reminding that the luminosity uncertainty for the no-pileup run is very large while the uncertainty of low μ run is 2.8%.

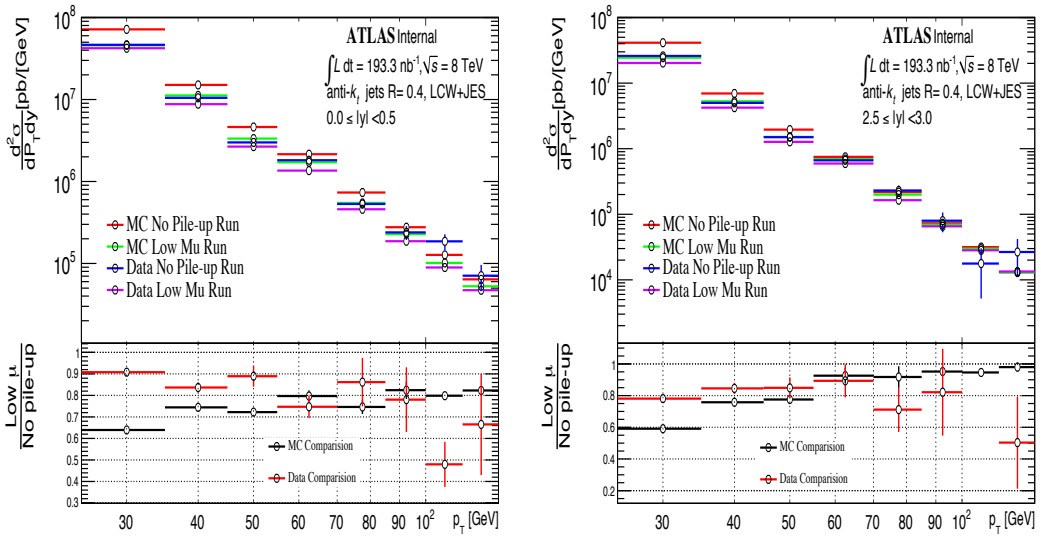


Figure 3.4: Comparison between no pile-up run and low μ run in MC and data using reconstructed jet p_T distribution for anti-k_t jets with R = 0.4. In the lower part of the figure the ratio between no pile-up run and low μ run in MC simulation(black line) and data(red line) is shown.

[9]JZXW, X=0,...8. These samples are generated with flat statistics in (leading) truth jet p_T cut

3.9 Unfolding

Detector effects will change the measurement of specific variables with respect to their "true" value. The measurements will be affected by acceptance, resolution, detector behaviour, calibration and so on. These effects can be included in the MC simulation and the data result can be corrected by accounting for them. The removal of detector effects from the measurement is known as unfolding, and it allows comparison of the result directly with theory or with other experiments. In the following, for a given variable the reconstructed distribution from the Monte Carlo can be represented as a histogram, projected into a vector that will be called R, while the truth-level distribution can be expressed by a vector called T. The vector holding the reconstructed distribution from data will be called D, and the unfolded distribution (from data, but after removing detector effects) will be U.

Several methods are used to unfold detector effects from experimental data. The bin-by-bin method consists in taking the ratio between the truth level and detector level distributions in the Monte Carlo $R = D/T$ for the observable; the unfolded distribution will be the bin-by-bin product of the data distribution times the inverse of that ratio: $U = DR^{-1}$. This method is correct only first order since it does not account for bin-by-bin migrations and relies on perfect description of data by MC. To overcome this limitations, iterative unfolding methods have been developed.

3.9.1 Principles of Iterative Unfolding

An obvious extension of the bin-by-bin method described above would be replacing the vector ratio R with a transfer matrix M whose components $M_{i,j}$ represent the probability that the true observable is in bin i and the reconstructed in bin j. So, by construction, $D = MT$; the unfolded distribution could be obtained by multiplying the data distribution by the inverse of M: $U = M^{-1}D$. The problem of this approach is that inverting large matrixes

can lead to large instabilities and fluctuations, especially in the case of limited statistics for the Monte Carlo used to build the transfer matrix. An alternative approach is the iterative Bayesian unfolding [69]. This method consists in making an initial hypothesis on the truth-level distribution of the quantity to be unfolded, that will be called vector T_0. This initial truth-level quantity gets multiplied by the transfer matrix M to obtain a first iteration of the reconstructed distribution: $R_0 = MT_0$. This first distribution is compared to the data histogram D, and the differences propagated back to the initial truth-level hypothesis, to obtain a second truth-level distribution T_1. This modified hypothesis is again multiplied by the transfer matrix, to obtain a second iteration of the reconstructed distribution, $R_1 = MT_1$, that is again compared to data. The procedure continues until a satisfactory agreement is obtained between the data distribution and the hypothetical detector-level one R_n. The corresponding truth-level distribution T_n is taken as the unfolded distribution U that will best reproduce the data after the application of detector effects.

This technique has been shown to be able to perform a good unfolding of complex distributions, but it can suffer from large fluctuations as the number of iterations increase. The iterative, dynamically stabilised (IDS) method [70] is a modification of the original iterative unfolding, but it makes use of a regularisation function, that reduces the amount of change in the successive truth-level distributions T_n as the number of iterations n increases; it also accounts for the fact that efficiency and purity correction derived by taking the ratio of spectrum from matched events (at reconstruction level for efficiency correction and particle level for purity correction) with that from all events. This is the method used in this analysis.

The MC simulation can be used to construct a transfer matrix that relates the reconstructed jet p_T at detector-level to the one at particle-level, called "truth jets" in the following, for a given jet rapidity bin. In the case of the inclusive-jet distributions, each MC jet provides an entry in the transfer matrix.

The matching between the truth and the reconstructed jet is performed

with the following algorithm:

- for each reconstructed jet the truth jet closest in ΔR is chosen.
- for each truth jet the reconstructed jet closest in ΔR is chosen.
- if the closest reconstructed jet to the truth jet i , is the reconstructed jet j, and at the same time, the closest truth jet to the reconstructed jet j is the truth jet i, and $\Delta R < 0.3$, the jets are considered "p_T-matched".
- If the truth jet and the reconstructed jet are p_T-matched and fall in the same rapidity bin, they are considered "matched".
- Truth jets that are not matched to a reconstructed jet are considered as "unmatched truth jets".
- Reconstructed jets that are not matched to a truth jet are considered as "unmatched reconstructed jets".
- if a truth jet is in a different rapidity bin than the p_T-matched reconstructed jet, the truth jet is considered "η-unmatched".
- if a reconstructed jet is in a different rapidity bin than the p_T-matched truth jet, the reconstructed jet is considered "η-unmatched".

The transfer matrix is filled with the matched jets only, and is shown in Figure 3.5 for anti-k_t jets with $R = 0.4$. Truth jets that are p_T-matched, but η-unmatched, are shown in the upper slice of the histogram shown below the transfer matrix. Reconstructed jets that are p_T-matched, but η-unmatched, are shown in the right slice of the histogram shown on the left to the transfer matrix. Finally, the p_T-unmatched truth jets are shown in the lower slice of the histogram below the transfer matrix. The p_T-unmatched reconstructed jets are shown in the left slice of the histogram on the left to the transfer matrix.

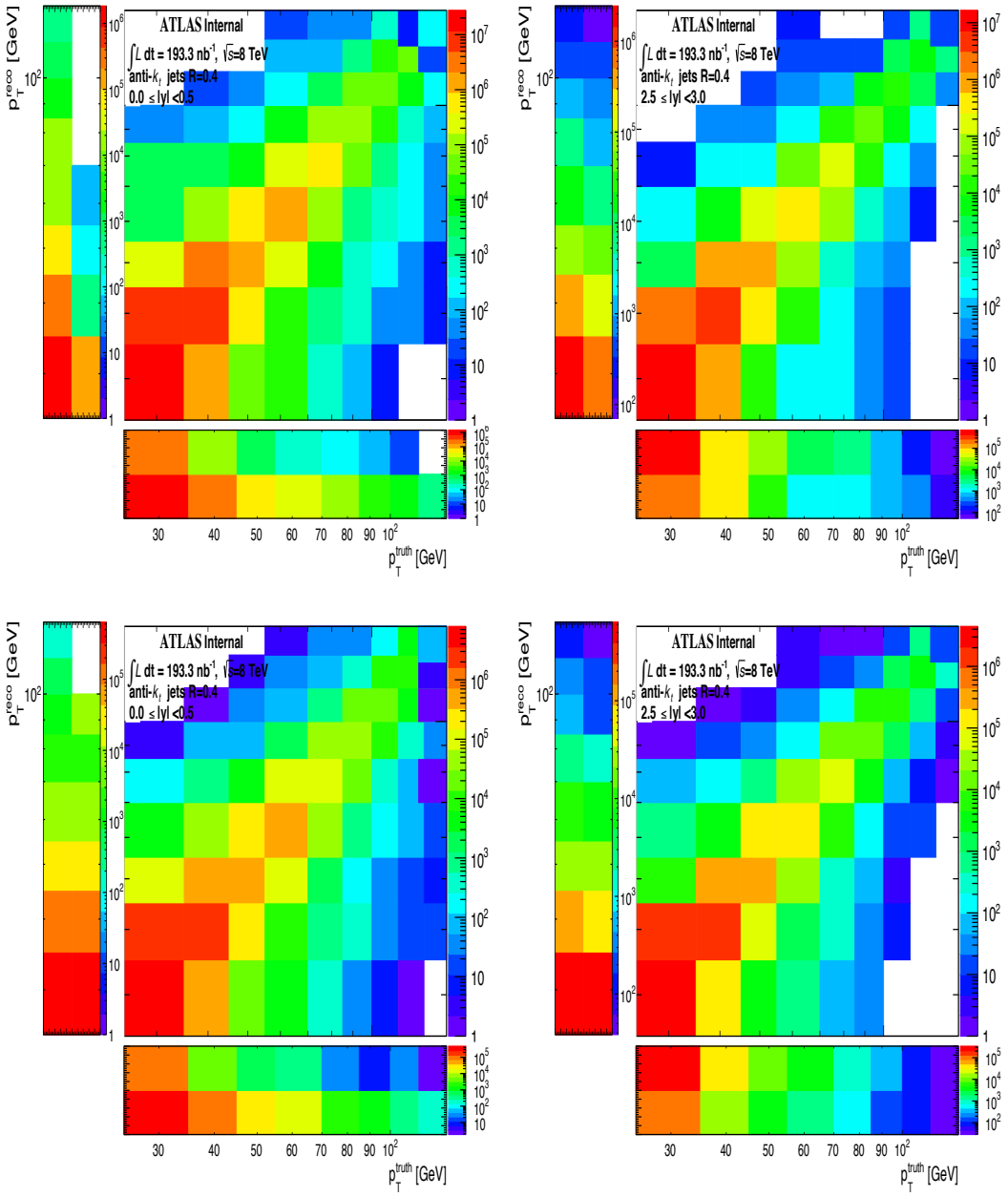


Figure 3.5: Transfer matrix relating the true to the reconstruction jet transverse momentum for anti-k_t R=0.4 at $0.0 \leq |y| < 0.5$ and $2.5 \leq |y| < 3.0$. The histograms to the left and below the transfer matrix show the number of unmatched jets. First two plots are for the no pile-up run (top) and last two ones (bottom) are for the low μ run.

3.9.2 Unfolding Process

There are three steps in the unfolding procedure for the correction of the data for resolution and efficiency effects.

In the first step, the matching efficiency is calculated taking the ratio of the reconstruction-level spectrum from matched events, obtained by a projection of the transfer matrix, with the reconstruction-level spectrum from all events, and the data spectrum is corrected for this efficiency. Figures 3.6 show the matching efficiencies for the no pile-up and low μ runs. Their dependence on transverse momentum is quite small, and as well as the absolute value of the inefficiency. At a second stage, the unfolding is performed using using the IDS method. As mentioned above the use of a transfer matrix accounts for bin-to-bin migrations; migration probabilities are given by

$$P_{i,j} = \frac{A_{i,j}}{\sum_k A_{i,k}} \tag{3.2}$$

for a jet generated in a given bin i to be reconstructed in bin j. The truth MC is re-weighted to the shape of the corrected data spectrum. Many iterations of the matrix application procedure can be applied to decrease sensitivity of the unfolded data to the shape of the initial hypothesis on the p_T distribution. However, for this particular analysis, one iteration is always sufficient to achieve convergence. The regularisation, preventing statistical fluctuations from being amplified by the successive iterations, uses the significance of the difference between the data and the reconstructed jets p_T spectrum in each bin.

Finally, the unfolded spectrum is corrected by a set of purity factors, derived by taking the ratio of the particle-level spectrum from matched events to that from all events.

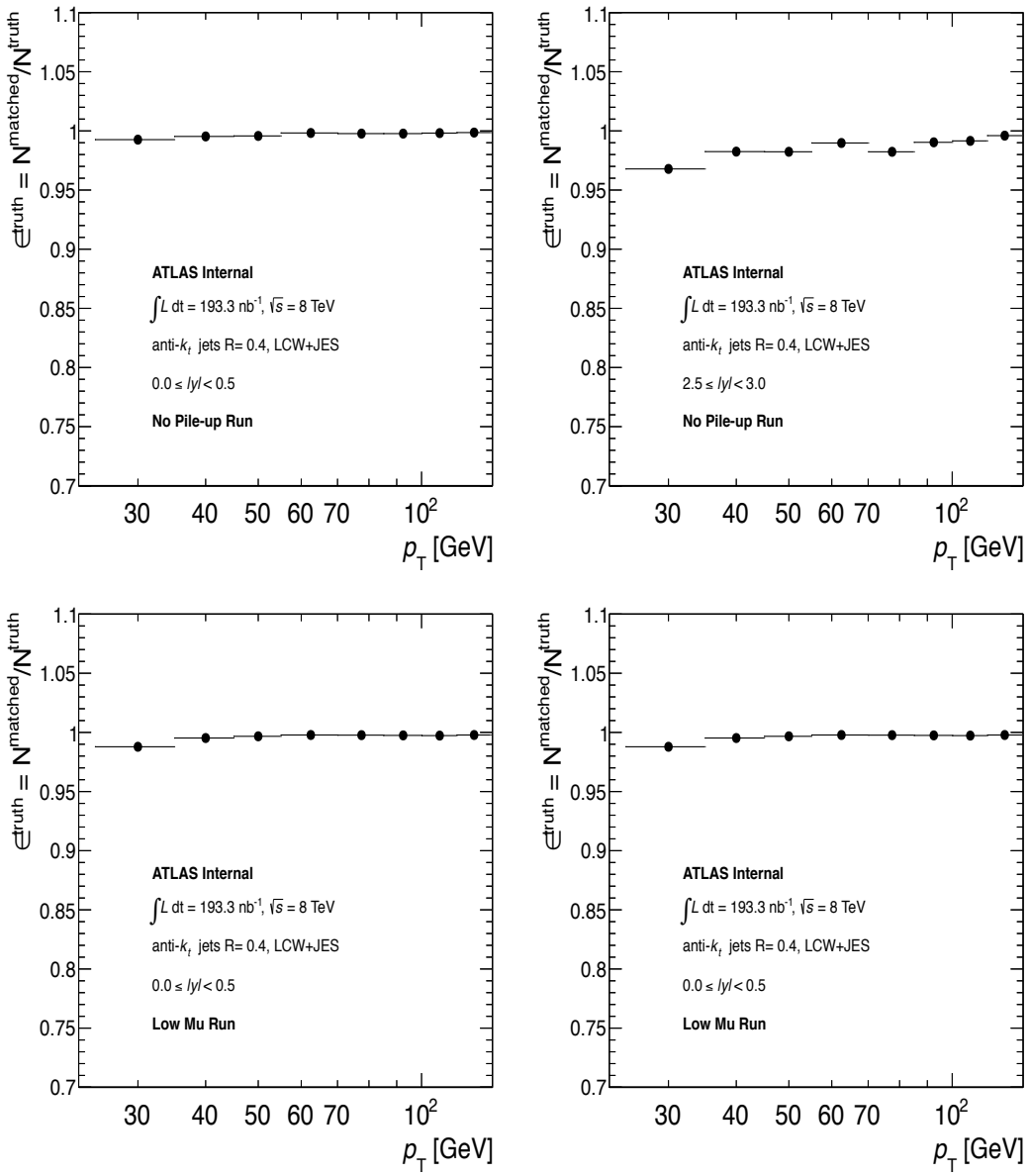


Figure 3.6: The fraction of p_T-matched truth jets with respect to total number of truth jets in given p_T bin for anti-k_t jets for $0.0 \leq |y| < 0.5$ (left) and right ones are $2.5 \leq |y| < 3.0$. First two plots (top) are for no pile-up and last two ones (bottom) are for low μ run.

3.9.3 Unfolded results for the no pile-up, low μ and high mu runs

Figure 3.7 shows a comparison of unfolded jet transverse momentum spectrum between the no-pileup and the low μ runs. A difference between them is still present after the unfolding, and is at the 20% level. It has to be recalled that the luminosity uncertainty of the no-pileup run at the 10% level, and could explain part of this discrepancy; jet calibration is also different due to the different calorimeter conditions.

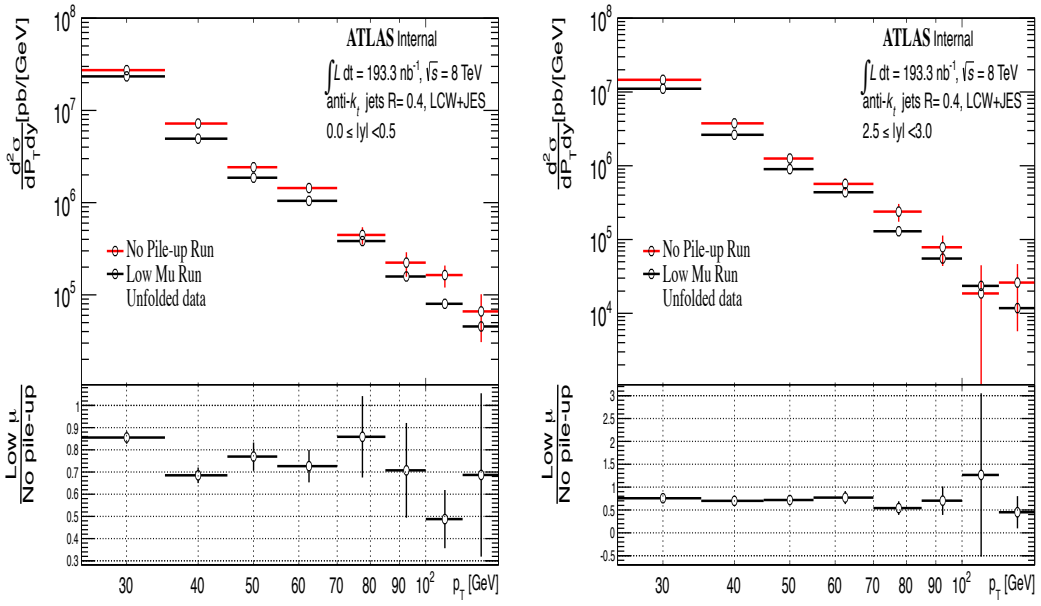


Figure 3.7: Comparision between no pile-up and low μ run in unfolded data at $0.0 \leq |y| < 0.5$ and right ones are $2.5 \leq |y| < 3.0$.

There is an overlap of three transverse momentum bins, in the region between 85 and 134 GeV. The comparison between the two measurements is shown in figure 3.8. The agreement is quite reasonable, considering that systematic uncertainties are not yet accounted for. In the Atlas publication, the boundary between them is set at 100 GeV, so low μ run results are used below that transverse momentum value.

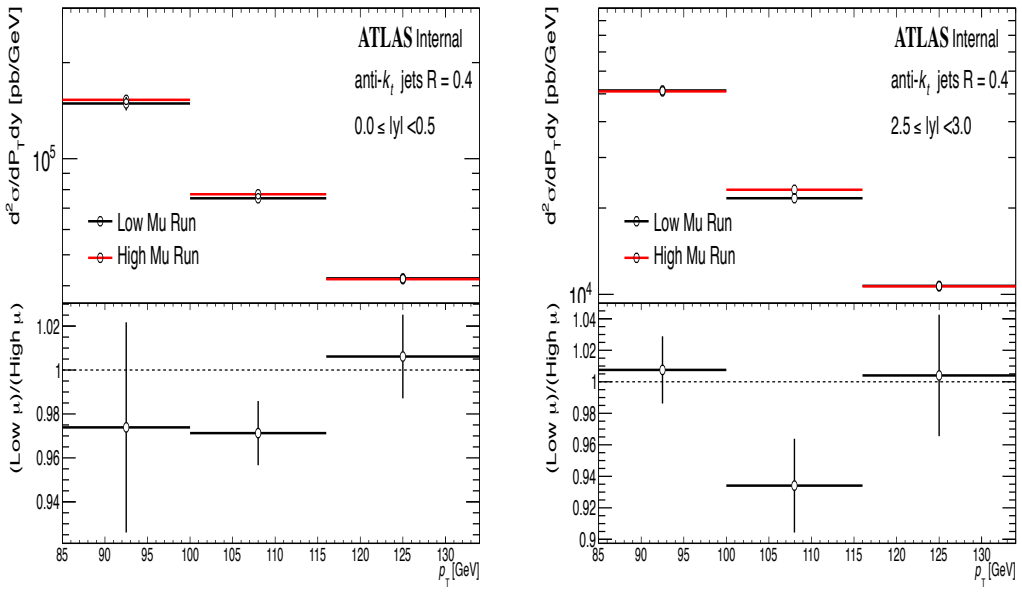


Figure 3.8: The plots represent comparison between low μ and high mu run at $0.0 \leq |y| < 0.5$ and right ones are $2.5 \leq |y| < 3.0$ in unfolded data.

3.10 Systematic and Statistical Uncertainties

The inclusive jet cross-section measurement is affected by the following main systematics effects:

- jet energy scale calibration (JES)
- jet energy resolution (JER)
- jet angular resolution (JAR)
- unfolding bias
- luminosity

JES uncertainty affects the measurement in the full phase space, since a change in transverse momentum scale will move jets from one bin to the other. If the jet energy resolution is different between data and Monte Carlo, the correspondence between truth and reconstructed jet transverse momenta

will be less precise, so migrations between bins is going to be affected. If jet angular resolution is badly described, migration between different rapidity bins will occur. All these uncertainties are propagated to the unfolded distribution by modifying the transfer matrix using Monte Carlo samples.

3.10.1 Uncertainty on Jet Energy Scale

The jet energy scale (JES) systematic uncertainty is the dominant one for measurement of inclusive-jet cross-section using low μ run [65]. 66 individual JES components are used that describe the correlations between the jet p_T and rapidity bins. A summary of all sources of jet energy scale uncertainties is given in Table 3.8.

The JES systematics is derived from in situ pT-balance techniques where the jet p_T is balanced against a well measured objects, i.e. a photon and Z-boson or a system of calibrated low-p_T jets [71] [72]. For the calibration in the central region Z+jet and γ+jet events are used up to about jet p_T of 1 TeV. For higher p_T the multi-jet balance technique is used where a high-p_T jet is balanced against a system of well calibrated low-p_T jets. For this highest-p_T, an uncertainty is derived from single particle response measurements in the ATLAS detector and in the test-beam. The transfer from the JES uncertainties in the central region to the forward region is performed using the dijet p_T balance technique where a jet in the central region balances a jet in the forward region (Dijet η-intercalibration) [72].

There are four components related to common sources on electron energy measurements and the photon in the Z+jet and γ+jet balance techniques. In the Z+jet balance, an important role is played by the electron energy scale. This is derived in situ using the Z-boson mass constraint. In addition, there are three systematic uncertainty components related to the transport of the electron energy scale uncertainty to the photon energy scale uncertainty; these are related to the dead material in front of the LAr calorimeter, the electromagnetic scale in the LAr pre-sampler and the electron energy calibration.

The suppression of energy induced by pile-up interactions plays an important role in the JES calibration of the 2012 data-sets. For the jet area correction techniques and for the residual jet energy off-set correction, two systematics components encode the residual dependence on the number of reconstructed vertices and the expected number of events $<\mu>$. The residual dependence on the jet p_T is also considered. In addition, one JES uncertainty component parameterises the difference in the event density between data and the Monte Carlo simulation [73].

The jet response depends on the fragmentation properties of the jets. In particular, there is a variation for quark- and gluon-induced jets of a few percent. Therefore, the fraction of quark- and gluon-induced jets in the event sample induces an uncertainty. Thanks to the use of the GSC calibration, the flavour response difference is reduced.

As shown in the previous inclusive jet measurements [74], the treatment of the detailed uncertainty components allows to consider in detail the correlations among the various sources. All the components of the JES uncertainty are propagated, using MC samples, through the unfolding procedure, taking into account their correlations (each component is assumed 100% correlated in p_T of the jets) and asymmetries. For each component, all the jet p_T values are shifted up and down by one standard deviation and the observable reconstructed; the two resulting spectra are unfolded, and the outputs are compared with the nominal unfolded spectrum. The difference between the output(up or down) and nominal unfolded spectrum gives the uncertainty for each component(up/down). A summary of all sources of jet energy scale uncertainties and the other common uncertainty sources; JAR,JER,unfolding discussed above is given in Table 3.8.

3.10.2 Uncertainty of Jet Energy Resolution

The fractional uncertainty on the jet p_{T} resolution is derived in situ using dijet p_{T} balance and the bisector method and through comparison with MC

Name	N	Description
Z+jet balance	1	
syst_JES_Zjet_MuScale	1	Muon momentum scale
syst_JES_Zjet_MuSmearID	1	Muon momentum resolution in inner detector
syst_JES_Zjet_MuSmearMS	1	Muon momentum resolution in muon detector
syst_JES_Zjet_Veto	1	Radiation suppression due to second jet cut
syst_JES_Zjet_dPhi	1	Variation of the fully balance event selection
syst_JES_Zjet_MC	1	MC generator
syst_JES_Zjet_KTerm	1	Contribution of soft particles outside the jet cone
syst_JES_Zjet_JVF	1	Jet selection from JVF pile-up jet suppression
syst_JES_Zjet_Stati	11	Statistical uncertainty for each of the 11 bins
γ+jet balance		
syst_JES_Gjet_Veto	1	Radiation suppression due to second jet cut
syst_JES_Gjet_dPhi	1	Variation of the fully balance event selection
syst_JES_Gjet_Generator	1	MC generator uncertainty
syst_JES_Gjet_OOC	1	Contribution of soft particles outside the jet cone
syst_JES_Gjet_Purity	1	Photon purity (background from multi-jet events)
syst_JES_Gjet_Stati	13	Statistical uncertainty for each of the 13 bins
Multi-jet balance		
syst_JES_MJB_Alpha	1	Angle between leading jet and recoil system
syst_JES_MJB_Beta	1	Angle between leading jet and closest sub-leading jet
syst_JES_MJB_Fragmentation	1	Jet fragmentation modelling uncertainty
syst_JES_MJB_Asym	1	p_T asymmetry selection between leading jet and sub-leading jet
syst_JES_MJB_Threshold	1	Jet p_T threshold
syst_JES_MJB_Stati	10	Statistical uncertainty for each of the 10 bins
Common sources in Z+jet and γ+jet balance		
syst_JES_LArEsmear	1	Photon and electron energy resolution in in situ balance
syst_JES_LArESmaterial	1	Photon energy scale due to dead material in front of LAr
syst_JES_LArESpresampler	1	Photon energy scale from pre-sampler calibration
syst_JES_LArESZee	1	Photon energy scale from electron energy scale in Z+jet balance
Dijet η-intercalibration		
syst_JES_EtaIntercalibration_Modelling	1	Generator modelling uncertainty in η-intercalibration
syst_JES_EtaIntercalibration_TotalStat	1	Statistical uncertainty
high-p_T jets syst_JES_SingleParticle_Highp_T	1	High-p_T calibration from single particle response
Pile-up substraction technique		
syst_JES_NPVOffset	1	Pile-up off-set substraction for primary vertices
syst_JES_MuOffset	1	Pile-up off-set substraction for number of expected interactions
syst_JES_Pileup_Pt_term	1	Pile-up jet area substraction technique p_T-term
syst_JES_Pileup_Rho_topology	1	Pile-up jet area substraction technique event density uncertainty
Common source in jet calibration	1	
syst_JES_Flavour_Comp	1	Quark/gluon jet composition difference various MC generators
syst_JES_Flavour_Response	1	Quark/gluon jet response difference between Pythia and Herwig
Common sources		
syst_jer	1	Jet resolution uncertainty
syst_jar	1	Jet Angular resolution uncertainty
syst_Unfolding_bias	1	Residual bias in the unfolding

Table 3.8: Summary table of the uncertainty components for each in-situ pT-balance technique and common uncertainty sources. Given is the name, the number of components and a short description. In case of JES terms, each entry corresponds to two components: one for the up and one of the down variation.

[75]. Dijet where the p_T of the leading and sub-leading jet are balanced are used to derive the resolution in the dijet balance technique. In the bisector method, the imbalance in the transverse momentum vector is defined as $\overrightarrow{P_T}$ $=\overrightarrow{P_T}^{jet,1}$ $+\overrightarrow{P_T}^{jet,2}$ using leading and sub-leading jets. The method considers its projections along an orthogonal coordinate system defined by the azimuthal

bisector $\Delta\phi$ between the two jets to derive the resolution.

In order to evaluate the systematic uncertainty due to a potentially worse resolution with respect to the Monte Carlo simulation, the nominal jet energy resolution is worsened by one standard deviation of its uncertainty. The p_T of each Monte Carlo jet is smeared by the factor σ_{smear}, calculated as:

$$\sigma^2_{\text{smear}} + \sigma^2_{\text{nominal}} = (\sigma_{\text{nominal}} + \Delta\sigma)^2 \tag{3.3}$$

where $\Delta\sigma$ is the uncertainty on the nominal fractional resolution,σ_{nominal}.

The relative difference of the data spectrum unfolded with the smeared transfer matrix and the data spectrum unfolded with the nominal transfer matrix (which is filled with the un-smeared MC sample) is taken as symmetrized systematic uncertainty. To reduce fluctuations due to limited statistics, each jet in the MC sample is smeared 1000 times using a different random draw through bootstrap method.

Adaptation of the JER uncertainty tool for the low μ and the no pile-run

With the new updated 2012 parameterisation of the JER, where the noise, stochastic response[10] and constant terms are fitted, the JER is applied to the low-mu and no pile-up conditions. In the new JER uncertainty tool, 9 up-/down- shift components are present. Calorimeter noise is well-simulated and does not contribute to the uncertainty for the high-μ run; however, an additional uncertainty is present for lower values of μ due to the extrapolation. As shown in Figure 3.9, the noise term is available for values of $<\mu>$ above 6; the extrapolation to lower values has been made using both a second-order and a third-order polynomial, and the difference between them is taken as an additional 10th component of the JER systematics.

The JER uncertainty is very small for the no pile-up run, it reaches up to

[10]stochastic response is a Poissonian event-to-event fluctuations with the constant term proportional to the jet energy (e.g. dead material)

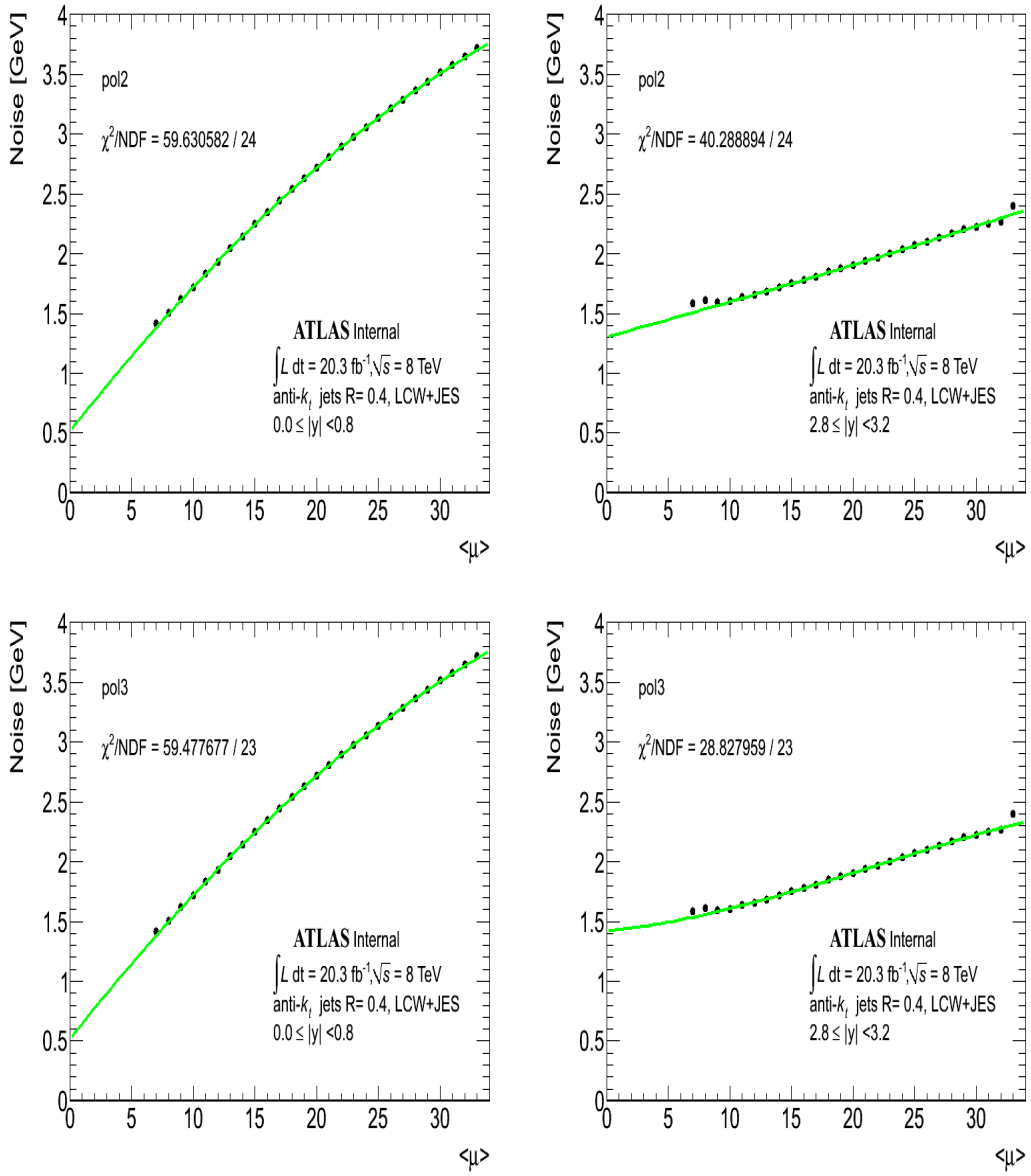


Figure 3.9: The noise term in the jet energy resolution fit as a function of the average number of additional pile-up interactions for anti-k_t jets with R = 0.4 at $0.0 \leq |y| < 0.5$ and right ones are $2.5 \leq |y| < 3.0$. Overlaid is a fit of a second order polynomial (above) and a three order polynomial (bottom).

8% for $p_T < 50$ GeV, and is at maximum 5% at higher p_T at central $|y|$ region for the low μ run as shown in Figure 3.14.

3.10.3 Jet angular resolution uncertainty

A systematic uncertainty on the jet direction must also be calculated, since a bad angular reconstruction could result in migrations between rapidity bins. The jet angular bias and resolution (JAR) in MC affects inclusive jet cross-section measurement [76]. The resolution is defined as a standard deviation of the difference in angles between truth and reconstructed jets. The same method is used as for JER uncertainty, namely comparisons between in-situ quantities between data and MC simulation.

The angular resolution on low pileup conditions is smaller at low energy than for the corresponding jets collected during high pileup conditions. As shown in Figure 3.14, the impact of this error is negligible at central $|y|$ region on no pile-run and low μ run. A maximum shift of 10% is found in the angular resolution at forward $|\eta|$ region.

3.10.4 Luminosity uncertainty

The final measured luminosity uncertainty is 2.8% on the 2012 data sample. This value refers to the high-μ as well as the intermediate-μ runs of this analysis. However, since the no-pileup run was taken during a period of very special accelerator conditions, the ATLAS luminosity group was not able to provide an estimation of luminosity uncertainty. From private communications, it was decided to use as a central value what would be obtained by considering standard conditions for the accelerator, but uncertainties around this value of up to 20% should be considered.

3.10.5 Statistical Uncertainty

In order to measure the statistical uncertainty on the cross section, the bootstrap method is used to take into account the fluctuations on the p_T spectrum

of reco jets on data, and those of truth and reconstructed jets in MC simulation. The first step in this technique is to produce a series of pseudo-experiments applying Poisson fluctuations to each event, and then build a set of replicas. The nominal transfer matrix is applied to each replica, and a set of fluctuated cross sections is generated by applying the transfer matrix to the fluctuated spectra; finally a covariance matrix is derived using the difference between each of the resulting unfolded spectra and their average. The full information on the magnitude of the statistical uncertainties coming from fluctuations in data and on the bin-to-bin correlations introduced by the unfolding procedure are stored in this covariance matrix. A second step accounts for the effect of limited Monte Carlo statistics. Each event in MC is filled into one of N replicas (1000 in this measurement) and generate N transfer matrixes. A second set of covariance matrices is created, and the unfolding performed with these transfer matrixes to account for the statistical fluctuations in the MC simulation. Figure 3.10 shows relative statistical uncertainty; Full errors are calculated by adding statistical fluctuations in data and MC simulation in quadrature on a bin-by-bin basis.

3.10.6 Estimating the systematic uncertainty due to shape differences between data and MC

A systematic uncertainty in the unfolding can arise from possible shape differences between data and the reconstructed MC spectrum. This comparison is used to build a closure test: for jets with true-reco matching, the truth MC is reweighted directly in the transfer matrix, by multiplying each column of the matrix by a given weight. These weights are chosen to improve the agreement between data (after applying the matching efficiency at reco level) and reconstructed MC (obtained by projecting the modified matrix onto the reco axis).

The modified reconstructed MC is unfolded using the original MC matrix, and the result is compared with the modified truth MC. The resulting bias is

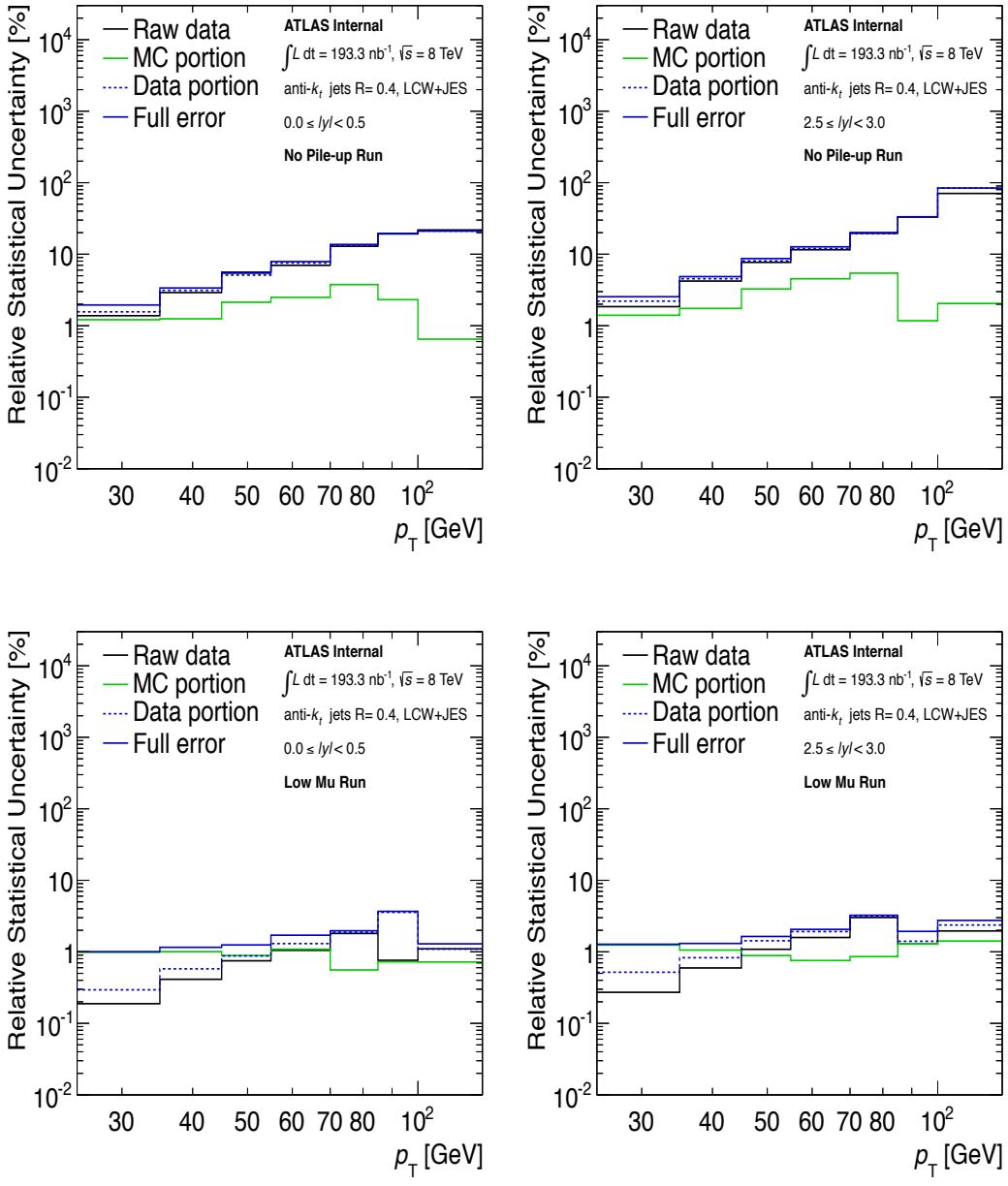


Figure 3.10: The relative statistical error is shown for no pile-up (above) and low μ run(bottom): for data before the unfolding (black line); for the MC simulation used for the transfer matrix (green line); for the data after the unfolding (dotted blue line), and for the combination of data and MC simulation after the unfolding (solid blue line)

interpreted as the systematic uncertainty, since it is defined as the difference between the unfolded result of the reweighted reconstructed MC (very similar to the data) using the original MC matrix, and the expected result for this unfolding test (i.e. the modified true MC). A series of toy (N) MCs events are generated weighting with a Poisson distribution , and each event is filled into one of N replicas, each representing a Poisson fluctuation of the original. The bias is calculated as the average of the biases obtained from the toy MCs.

Figure 3.11 shows the ratio of the data with inefficiency applied to the matched reconstructed Monte Carlo for the no pile-up and for the low μ runs, for two rapidity bins. Because the main goal here is to compare the shapes of the data and the MC at high pT, the curve for the total reweighted reconstructed MC has been normalised to that for data in the second p_T bin in each y bin. The reweighted Monte Carlo jet p_T distribution describes the data well. Without reweighting, shape differences up to 0.3% are observed at central $|y|$ region, growing to 4% in the forward region. The reweighted MC reproduces the curvature of the data spectrum at high pT, so the closure test is a reliable method to estimate the shape uncertainty. Also shown are the jet p_T distributions of the unmatched truth and reconstructed jets normalised to the original MC distribution. The fraction of unmatched jets is very small for all p_T bins, as expected.

Figure 3.12 shows the relative biases (i.e. the relative difference between the unfolded and the truth jet cross sections) obtained from the MC simulation only for the IDS method without iterations and with one iteration, and from the bin-to-bin unfolding method. It can be seen that while no-iteration IDS and the bin-to-bin method exhibit quite large biases in some cases, the bias from the IDS methods after one iteration is less than a fraction of %. These small deviations are included in the unfolding systematics uncertainty. The IDS method with one iteration has been chosen for this measurement.

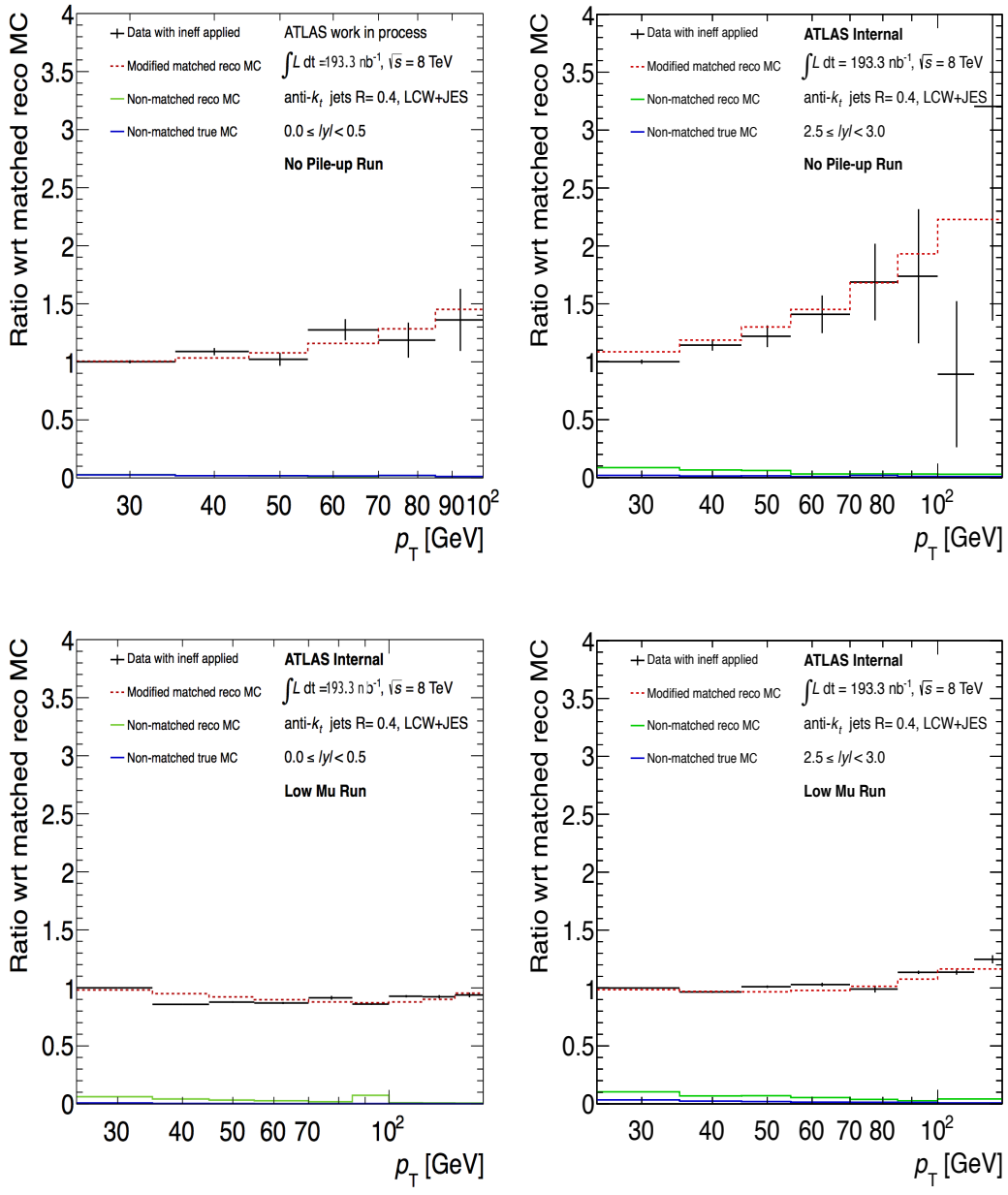


Figure 3.11: A comparison of the matched portion of the MC simulation with the efficiency corrected data, for the central (left) and forward (right) rapidity regions, for the no-pileup (top) and low μ (bottom) runs. Data are the black markers, modified MC simulation after the fit to a second-order polynomial are the dotted red line, while the unmatched reconstruction-level MC simulation are in green and the unmatched particle-level MC simulation in blue.

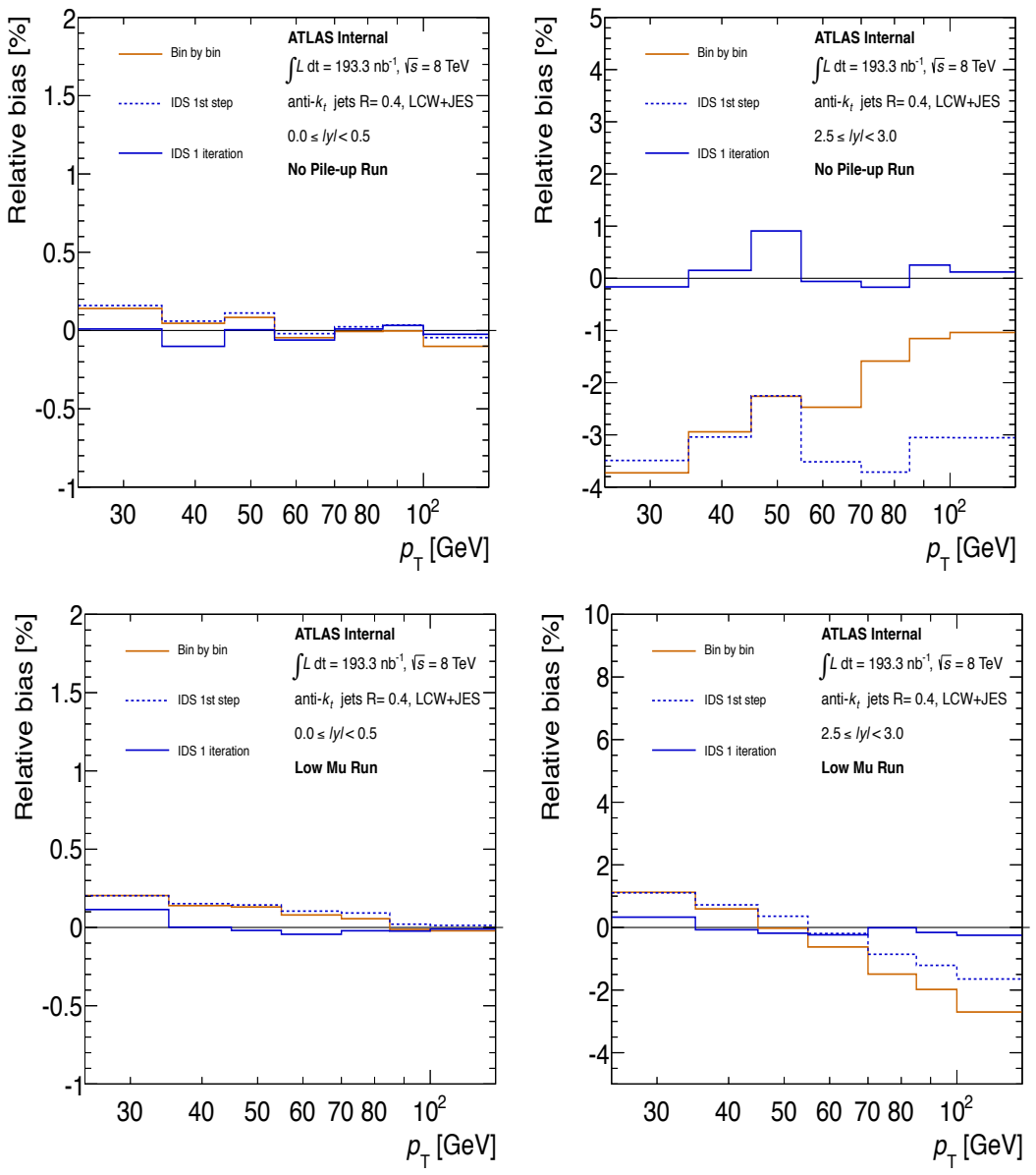


Figure 3.12: Relative bias due to the Monte Carlo shape differences for several unfolding methods. Top row refers to the no-pileup run, bottom to the low μ run. Left plots are for the central rapidity region, right plots are for forward rapidity.

3.10.7 Statistical Fluctuations and Smoothing

To reduce statistical fluctuations in the systematics uncertainties, a smoothing technique is applied. The procedure followed is described in detail in Ref. [77].

For each systematic uncertainty, the relative statistical uncertainty is computed using the bootstrap method (see [78] and [79]) in the detector unfolding as mentioned before. Using the statistical uncertainties the bins for each systematic uncertainty component are combined iteratively from both the left and right sides of the spectra, until their significance is $> 2\sigma$. The combination with the most bins remaining is chosen. Then a Gaussian kernel is used to fill the fine binning used of the measurement and to smooth out any additional statistical fluctuations. One bin below the reported range is used so that shape information is not underestimated for the first p_{T} bin, and the value of the largest bin is kept as the result after only the first step (rebinning by significance) since there are no bins above.

Examples of this smoothing can be seen in Figure 3.13, where a selected range of shapes and magnitudes are shown. The result is a reduction of statistical fluctuations in the systematic uncertainty components with minimal added bias.

The procedure also has the effect of reducing artificial asymmetries, which are shown to cause problems when performing statistical tests for agreement between data and theory (see [77, Section 11.2.2]).

As a summary, figures 3.14 show the relative total systematic uncertainty, along with the separate uncertainties on jet energy scale (JES+1σ and JES-1σ), the jet energy resolution, JAR, the unfolding bias and the statistical uncertainty for each jet rapidity bin.

The uncertainty on JES is the dominant uncertainty as expected. The total uncertainty increases up to 50% in the forward region for the no pile-up run, and up to 38% in the forward region for the low μ run.

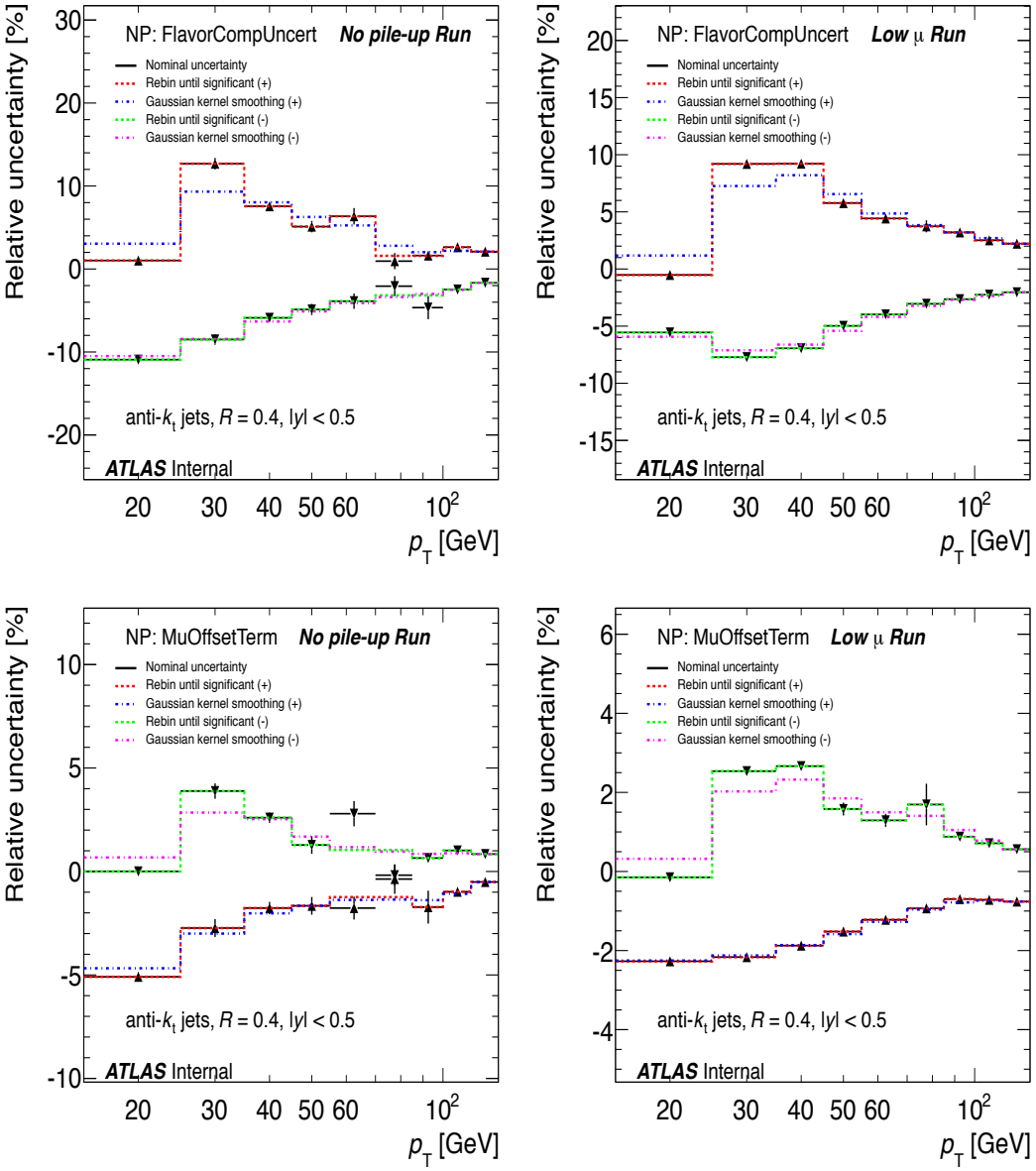


Figure 3.13: Relative uncertainty of the inclusive jet cross section measurements as a function of the jet p_T for the $0.0 \leq |y| < 0.5$ rapidity bin and for anti-k_t jets with R = 0.4. The following sources of JES uncertainty are shown: Flavor, Mu OffSet Term. Shown is the nominal uncertainties as obtained from the unfolding (crosses) and the smoothed one obtained after rebinning the bins that are not statistically significantly different (red dashed line), and the ones obtained after Gaussian kernel smoothing (dashed blue line). Plots on the left column refer to the no pile-up run, those from the right to the low μ run.

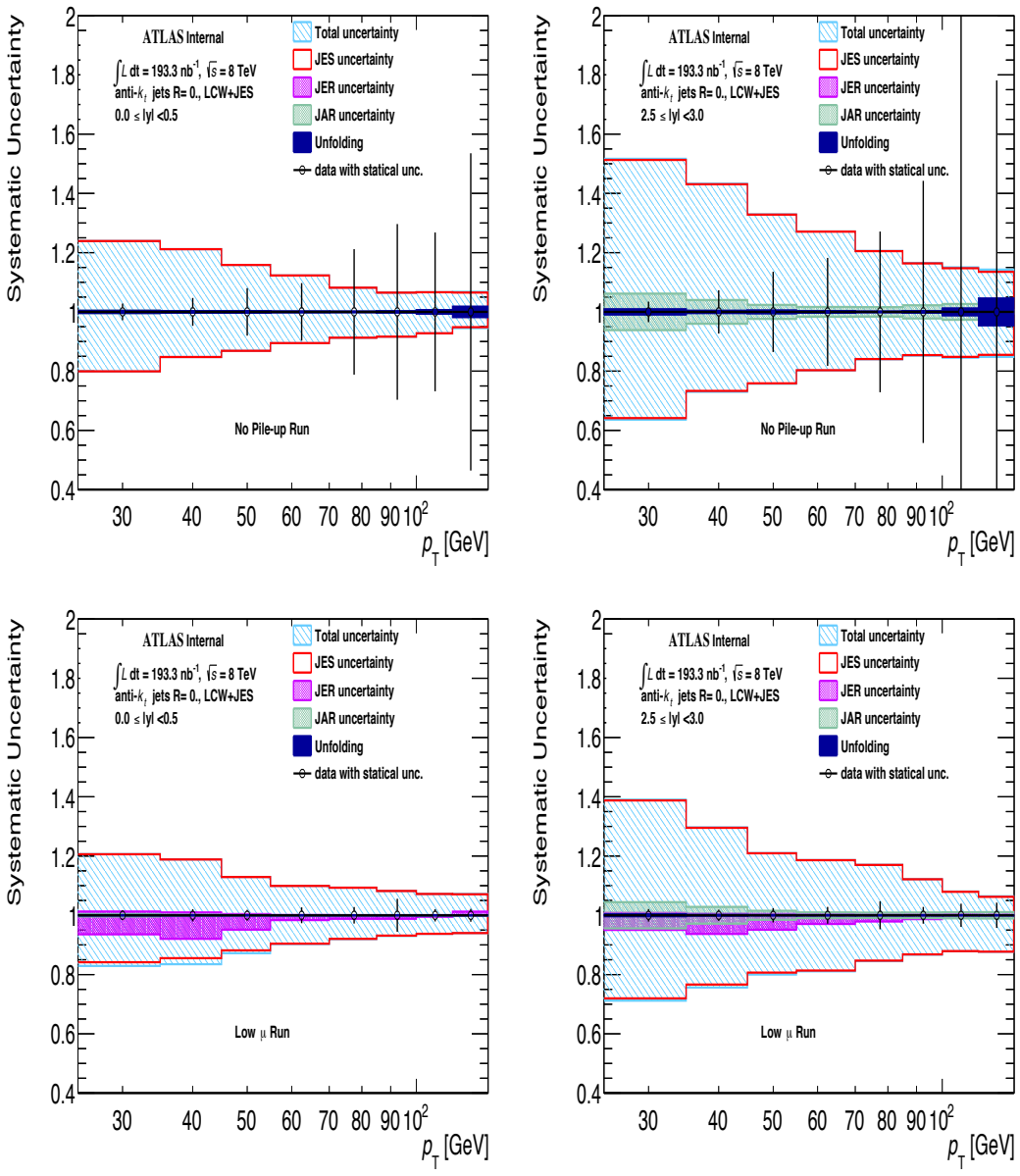


Figure 3.14: Relative systematic uncertainty for the inclusive jet cross-section as a function of p_T for the no pile-up (above) and the low μ run (bottom). The blue line represents the total uncertainty, the red is JES, purple is the JER, green JAR and blue is the unfolding bias. Shown are up- and down-variations.

3.11 Theory uncertainties

Theory uncertainties are evaluated using several PDF sets, changing the factorisation and renormalisation (R+F) scales, and the value of the strong coupling α_s constant according to its world-average uncertainty.

Renormalisation and factorisation scales: These quantities are connected to the QCD calculations used for the theory predictions. In this measurement, they are different on an event-by-event basis, and the default value for both is taken as the transverse momentum of the leading jet:

$$\mu_R = \mu_F = p_T^{max} \tag{3.4}$$

where μ_R is the renormalisation scale, μ_F the factorisation scale, and $p_T^{max}(y_i)$ is the maximum jet transverse momentum in the event. To estimate this uncertainty, both contributions are varied by a factor of two (up and down from its nominal value), considering all possible permutations except the opposite directions variation, which can bring instabilities due to large logarithm ($\log \frac{\mu_R}{\mu_F}$). The uncertainty is quoted as the envelope of all variations, i.e. the maximal deviation between the varied to the nominal scale setting.

PDF uncertainty: For most PDF sets, the uncertainties on the PDF fit parameters are expressed in terms of an orthogonal basis, called "Eigenvectors". Each eigenvector is associated with an eigenvalue, with a positive and a negative error. To calculate the total uncertainty due to PDF's, each of the eigenvalues has been varied by its positive and negative uncertainty, and the cross-section recalculated. The envelope of all variations is calculated as the total uncertainty. Four PDF sets have been considered in this study: CT10nlo [19], HERAPDF15NLO_EIG [20], NNPDF 3.0 [21] and MSTW2008-nlo68cl.

Strong coupling constant: The uncertainties on the value of the strong coupling constant are calculated by changing the value of α_s for each PDF sets (each PDF set uses a slightly different value of the coupling constant,

also different from the world average). The nominal value of cross-section has been changed in the two directions, α_s (down) and α_s(up), and uncertainty is calculated as:

$$\Delta\sigma_{up} = \frac{\sigma(\alpha_s(up)) - \sigma(\alpha_s(nominal))}{\alpha_s(up) - \alpha_s(nominal)} \times \Delta\alpha_s, \tag{3.5}$$

$$\Delta\sigma_{down} = \frac{\sigma(\alpha_s(nominal)) - \sigma(\alpha_s(down))}{\alpha_s(nominal) - \alpha_s(down)} \times \Delta\alpha_s, \tag{3.6}$$

where $\sigma(\alpha_s(up))$ and $\sigma(\alpha_s(down))$ are the inclusive cross sections calculated using a specific PDF set with the α_s up and down variations and $\Delta\alpha_s$ is the uncertainty on the nominal as calculated as:

$$\Delta\alpha_s = \frac{\Delta\alpha_s(WA)}{\alpha_s(WA)} \times \alpha_s(nominal), \tag{3.7}$$

where α_s(WA) is the world average α_s value (α_s(WA) = 0.1184) and $\Delta\alpha_s$(WA) its uncertainty. The world average uncertainty is assumed to be $\Delta\alpha_s$(WA) = 0.0012 as recommended in Ref. [80].

Summary of theory uncertainties Figures 3.15 and 3.16 show the effect of theory uncertainties, for the different PDF sets. Overall, the uncertainties due to the various PDF sets are similar to each other. The largest uncertainty corresponds to the change of R+F scales, and decreases slowly as a function of p_T, from $^{+2.8}_{-13}$% at low p_T to $^{+4}_{-8}$% at higher p_T, for the central region. The total uncertainty in the forward region does not change too much with respect to the central one.

3.11.1 Non-perturbative corrections

To compare the parton level NLO QCD calculations to data, their results need to be corrected for non-perturbative effects. The theory prediction for the inclusive jet cross-section at particle level can be written as:

$$\sigma_{theory} = \sigma_{\mathrm{NLO}} * k_{NP}, \quad \text{with} \quad k_{NP} = \frac{\sigma_{\mathrm{hadrons}}}{\sigma_{\mathrm{partons}}} \tag{3.8}$$

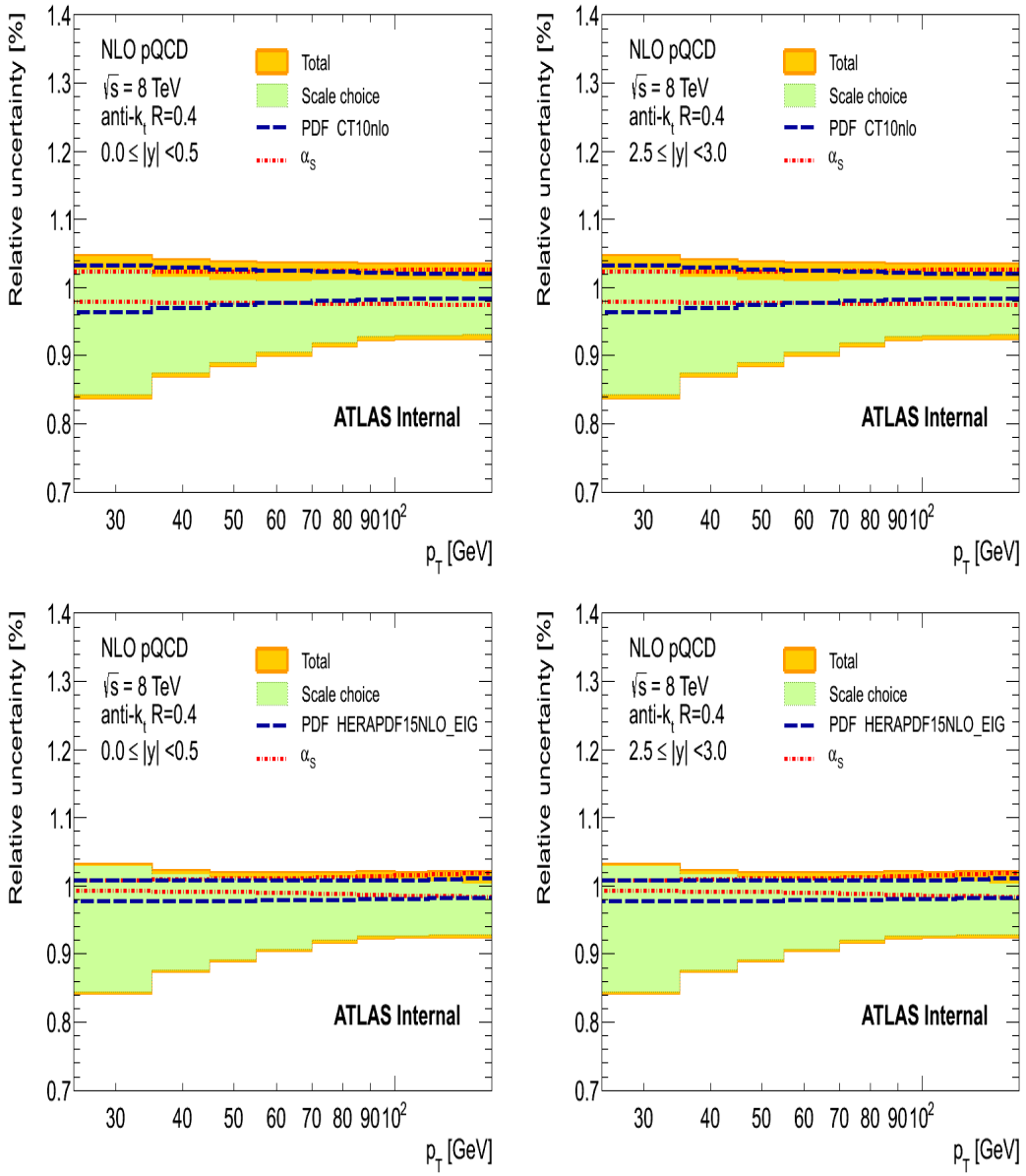


Figure 3.15: Relative theory uncertainties as a function of transverse momentum for $0.0 \leq |y| < 0.5$ (left) and $2.5 \leq |y| < 3.0$ (right), for the CT10nlo (top) and HERAPDF15NLO (bottom) PDF sets. The various components are R+F scales (blue), PDFs eigenvalues (black), α_s (red) and overall (orange).

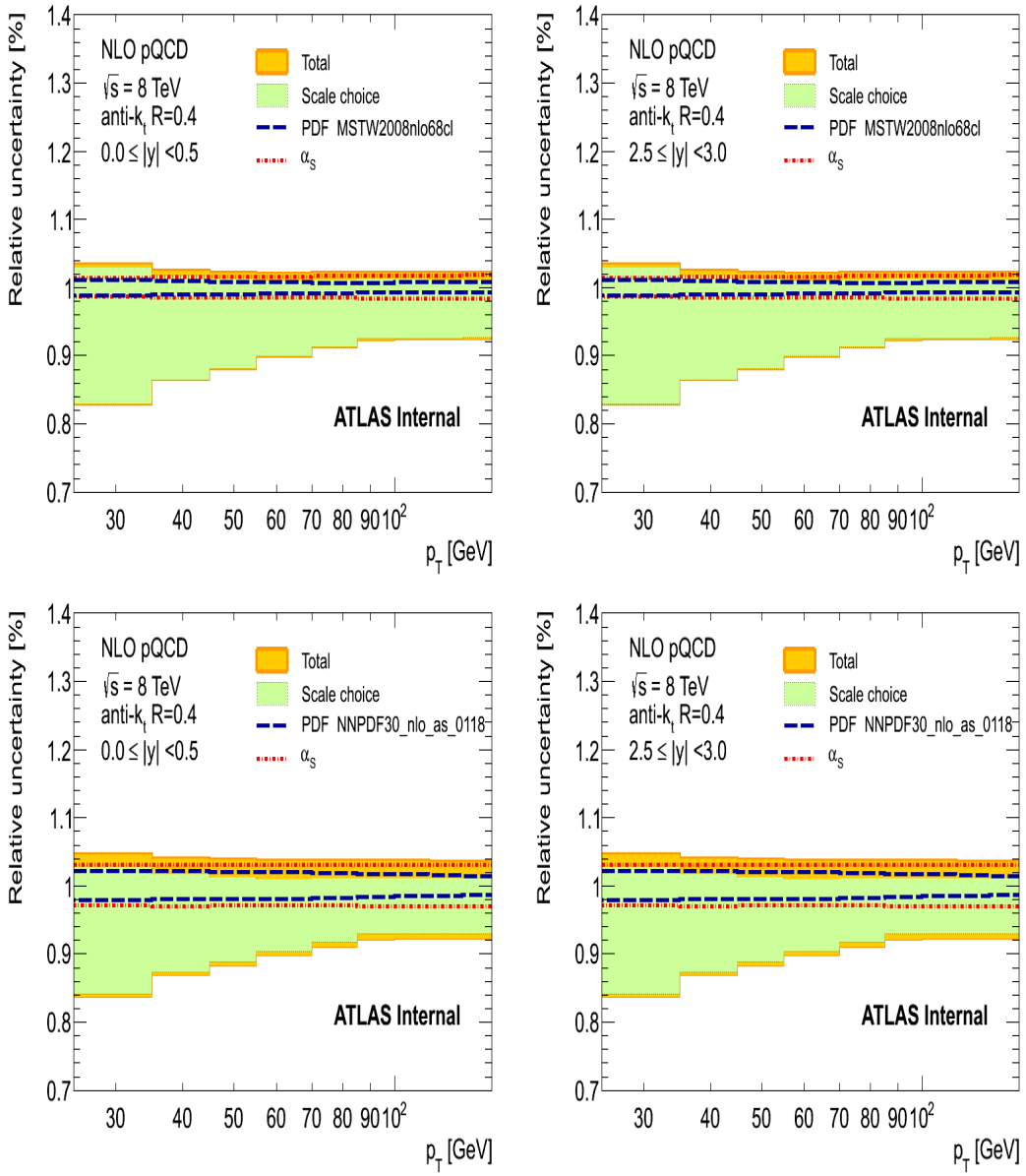


Figure 3.16: Relative theory uncertainties as a function of transverse momentum for $0.0 \leq |y| < 0.5$ (left) and $2.5 \leq |y| < 3.0$ (right), for the MSTW2008nlo68cl (top) and NNPDF 3.0 (bottom) PDF sets. The various components are R+F scales (blue), PDFs eigenvalues (black), α_s (red) and overall (orange)

where the symbol σ_{NLO} denotes the parton-level calculation (see Section 3.11). In order to calculate the correction k_{NP}, the inclusive cross-sections at different $|y|$ regions for anti-k_t jets with R=0.4 and 0.6 are produced using LO generators, with and without hadronisation and underlying event, giving the terms σ_{hadrons} and σ_{partons}. Hadron-level cross sections are built from the stable particles of the event record, and parton-level ones from final-state partons. To calculate the non-perturbative corrections, the bin-by-bin ratio (k_{NP}) between the cross-section measurement with and without hadronisation and the underlying event are computed using the generators and tunes shown in Table 3.9. More information about Pythia 6 tune can be found in Ref. [81], about the Pythia8 tune in Ref. [82], and about the Herwig++ tune in Ref. [83].

Name	Generator	Tune	PDF
pythia6_P2012	pythia 6	P2012	
pythia6_P2011C	Pyhtia6	P2011C	
pythia6_AUET2BLO	Pyhtia6	AUET2B	LO
pythia6_AMBT2BCTEQ	Pyhtia6	AUET2B	CTEQ
pythia6_AUET2BCT10	Pyhtia6	AUET2B	CT10
pythia8_4C	Pyhtia8	4C	
pythia8_AU2CT10	Pyhtia8	AU2	CTEQ
herwigpp_CTEQEE4	Herwigpp	CTEQEE4	

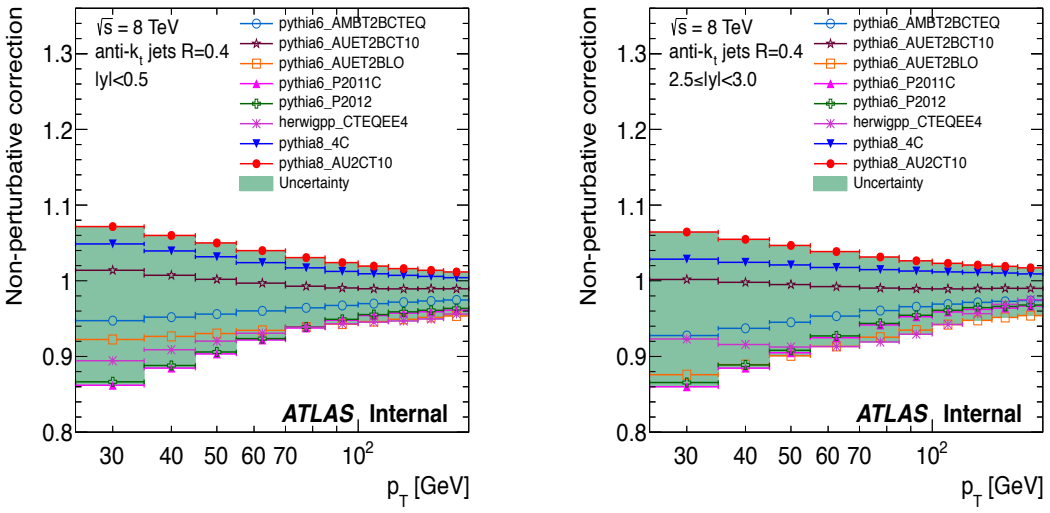
Table 3.9: Summary of MC tunes used for the evaluation of the non-perturbative corrections. The name of the generator, the soft physics model tune as well as the PDF set used for the hard process is specified.

The non-perturbative corrections are fitted with an exponential function to avoid statistical fluctuations :

$$\text{NPC}(p_{\text{T}}) = a + b\ \exp\left(c\, p_{\text{T}}\right) + d\, x, \tag{3.9}$$

where (a, b, c, d) are the parameters of the fit. The results of the fits for the various tunes are shown in Figure 3.17.

The NLOJet++ prediction is then corrected on a bin-by-bin basis by the central value of the correction, and the envelope of these ratios for all the tunes is taken as the total uncertainty.



(a) $0.0 \leq |y| < 0.5$ (b) $2.5 \leq |y| < 3.0$

Figure 3.17: Non-perturbative correction factors as a function of jet p_T for two ranges of $|y|$ at $0.0 \leq |y| < 0.5$ and $2.5 \leq |y| < 3.0$, shown for jets defined by the anti-k_t algorithm with R = 0.4. The corrections are derived using Pythia 6, Pythia 8 and Herwig++ with several soft physics tunes.

3.12 Electroweak Correction

The fixed order NLO QCD is corrected for electroweak effects as computed in Ref. [84]. The corrections include the effects of γ and $W^{\pm}/Z$ interactions at tree- and one-loop-level and are derived applying a NLO calculation for electroweak processes to a LO QCD calculation. In general the corrections to the cross-section are small even for large y bins. In the high-p_T tail of the inclusive jet spectra, radiative electroweak correction starts to be important. For the inclusive cross-section, this electroweak correction starts playing a role for values of p_T above 134 GeV, so it has been ignored since it is negligible in the p_T range relevant for this measurement.

Chapter 4

Result on the Measurement of inclusive jet cross-section

The inclusive double-differential jet cross-section as a function of jet p_T and $|y|$ is shown in Figures 4.1 for jets reconstructed with R=0.4 and 4.2 for R=0.6, in the kinematic region of the jet transverse momentum $p_T > 25$ GeV and $|y| < 3$ in steps of 0.5 rapidity units. In these Figure, no pile-up and low μ run data are compared with the NLO QCD theory calculations by NLOJet++ convoluted with the CT10nlo PDF set. Theory calculations also account for the non-perturbative correction. The cross sections are steeply decreasing as a function of p_T, spanning approximately 3 orders of magnitude. In general, no pile-up data has a larger cross-section than the theory prediction and of the low-μ run data over the full kinematic range for R=0.4. It has to be reminded that a large luminosity uncertainty, of the order of 20%, is present for that run. The absolute cross sections using jets with radius parameter R = 0.6 are higher than those obtained using R = 0.4 because the larger value of the jet radius parameter leads to more contributions from the parton shower and the underlying event. As mentioned before, the aim of this study is to merge the results from the low-μ and the high μ runs at the bin between 100-116 GeV, to obtain an overall measurement spanning a large range in jet transverse momentum. Due to the problems with luminosity determination, it was decided not to use the no-pileup run, but only the low-μ one, in the

ATLAS publication; however, both periods are used in this thesis to provide a final combination. This allows to perform a measurement of the inclusive cross-section, starting from a minimum transverse momentum of 25 GeV. Figures 4.3 and 4.4 show the merged inclusive jet double-differential cross-section as a function of p_T combining the no-pile run between 25-45GeV, the low μ run between 45-100GeV and the higher μ runs at $p_T > 100$GeV. The measurement is compared to a NLO QCD prediction using CT10nlo PDF set corrected for non-perturbative effects. This calculation gives an overall good description of the data. In the central part of the detector the cross-section in data falls by 8 order of magnitudes from 10^{11} pb^{-1} to 10^{3} pb^{-1} in the most central rapidity region $|y| < 0.5$ and from 10^{6} pb^{-1} to 10^{-5} pb^{-1} in the most forward 2.5 $\leq |y| <$3.0 at R=0.4. Jet transverse momenta of up to p_T= 2 TeV are reached for the first three $|y|$ bins. In the most forward region, the jet p_T reaches about 1000 GeV.

4.1 Ratios between theory and data

The ratios of the NLO pQCD predictions (including non-perturbative corrections) to the measured cross-sections are shown in figures 4.5- 4.8 for R=0.4 and 4.9- 4.12 for R=0.6. In these figure, no pile-up and low μ run data are considered separately, and compared to NLO predictions convoluted respectively with the PDF sets CT10, HERAPDF1.5, MSTW 2008, NNPDF 3.0. In each figure, only one PDF set is considered. It can be observed that in some distributions some discontinuity is present for the low-μ run after the first bin, in correspondence to 35 GeV. This is due to the fact that the jet energy resolution in that region is larger than the bin size, therefore the first bin still suffers from boundary effects in the unfolding. The results from all PDF sets for R=0.4 are higher than of low μ run for $p_T > 35$ GeV but this discrepancy is not observed for R=0.6. The difference between the no pile-up run and the theory predictions is much smaller. Finally, in figures 4.13 and

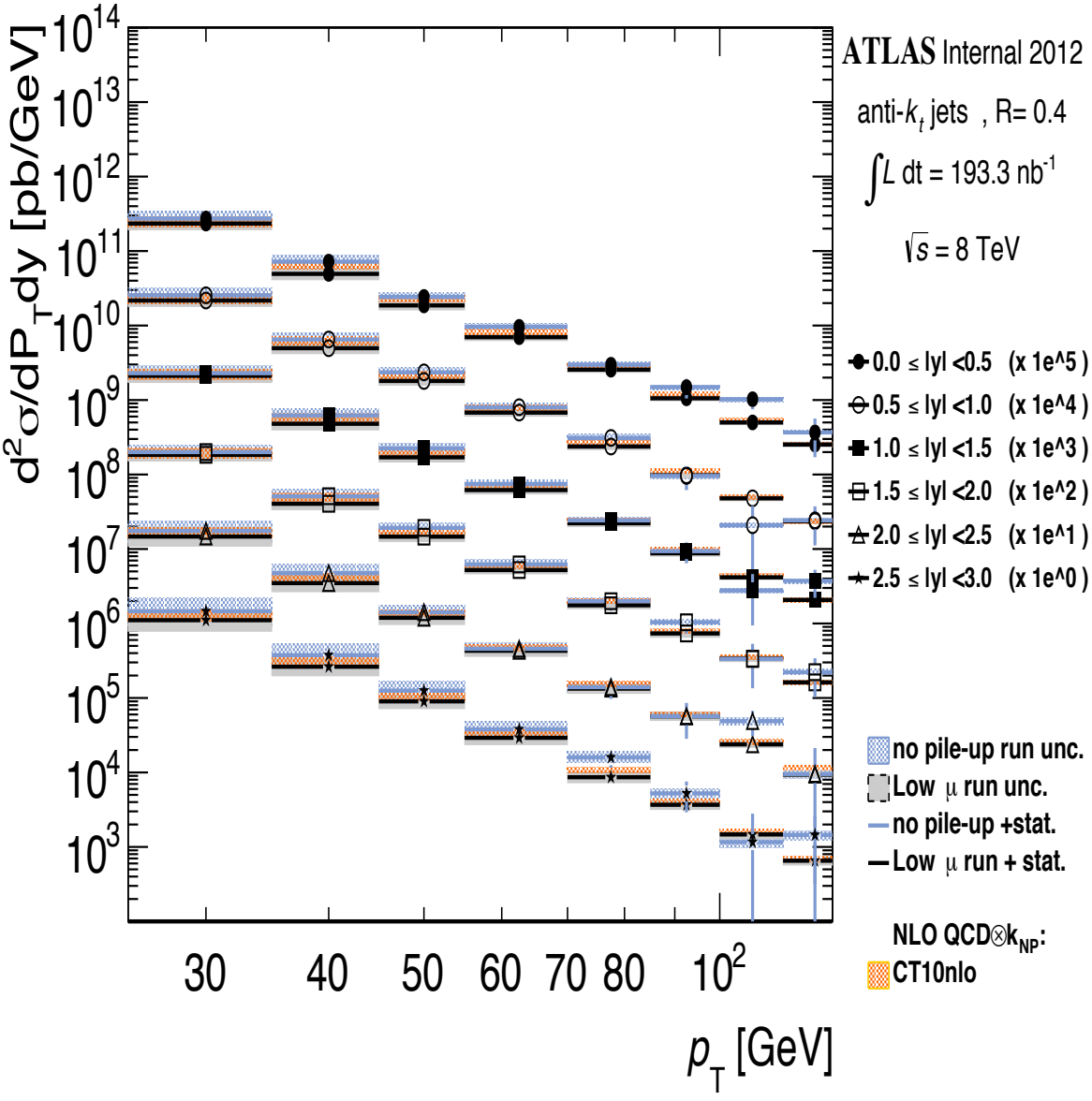


Figure 4.1: Inclusive jet double-differential cross sections for anti-k_t jets with radius parameter R = 0.4, shown as a function of p_T and $|y|$. To help visibility, the cross sections are multiplied by the factors indicated in the legend. The blue error bars indicates the statistical uncertainty of the no pile-up measurement, the black line the low μ run. The blue-shaded band indicates the sum in quadrature of the experimental systematic uncertainties on no pile-up run and the grey band is for low μ run data. For comparison, the NLO QCD predictions of NLOJet++ using the CT10 PDF set, corrected for non-perturbative, are included. The orange hatched band shows the uncertainty associated to the theory predictions.

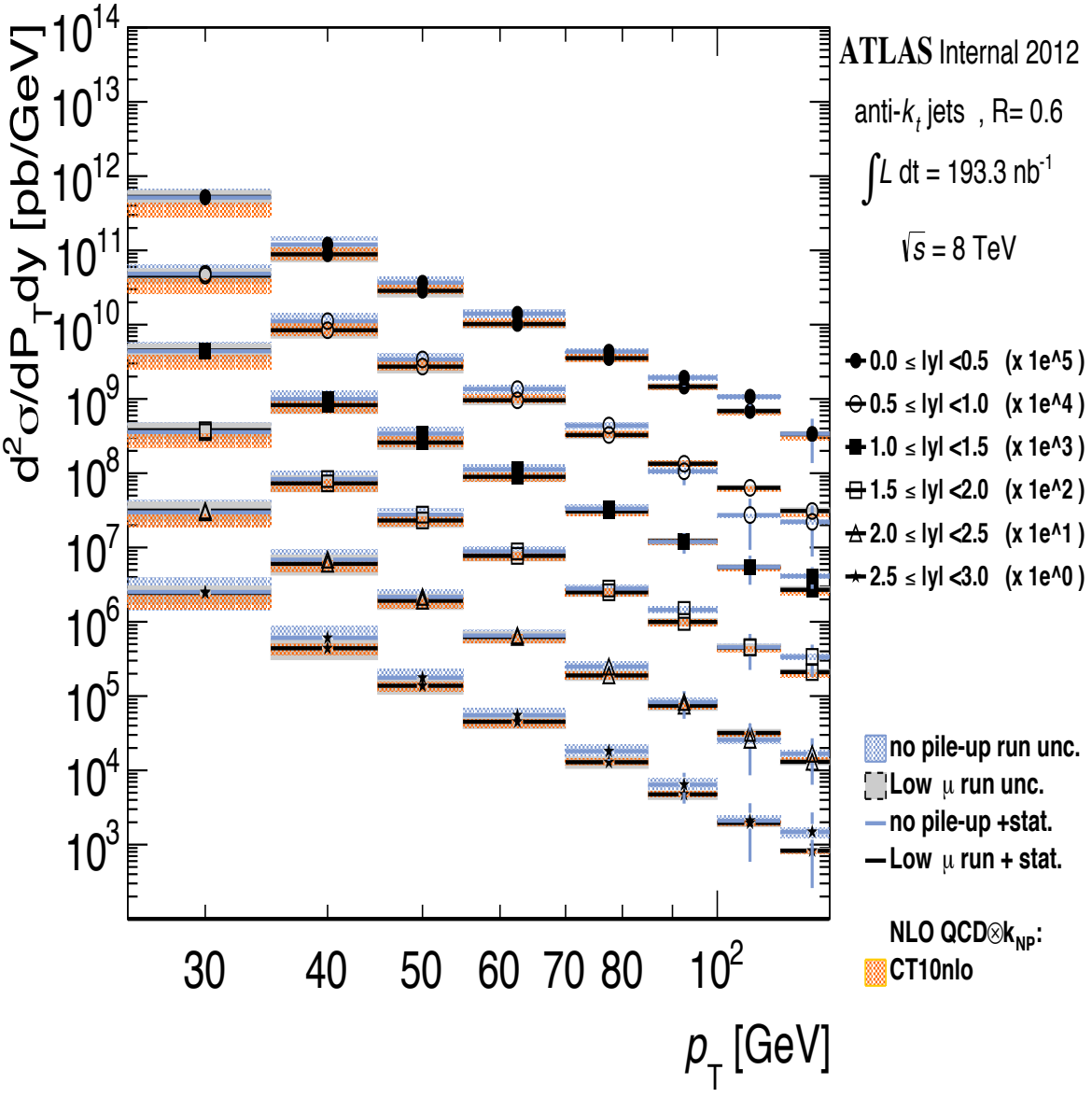


Figure 4.2: Inclusive jet double-differential cross sections for anti-k_t jets with radius parameter R = 0.6, shown as a function of p_T and $|y|$. To help visibility, the cross sections are multiplied by the factors indicated in the legend. The blue error bars indicates the statistical uncertainty of the no pile-up measurement, the black line the low μ run. The blue-shaded band indicates the sum in quadrature of the experimental systematic uncertainties on no pile-up run and the grey band is for low μ run data. For comparison, the NLO QCD predictions of NLOJet++ using the CT10 PDF set, corrected for non-perturbative, are included. The orange hatched band shows the uncertainty associated to the theory predictions.

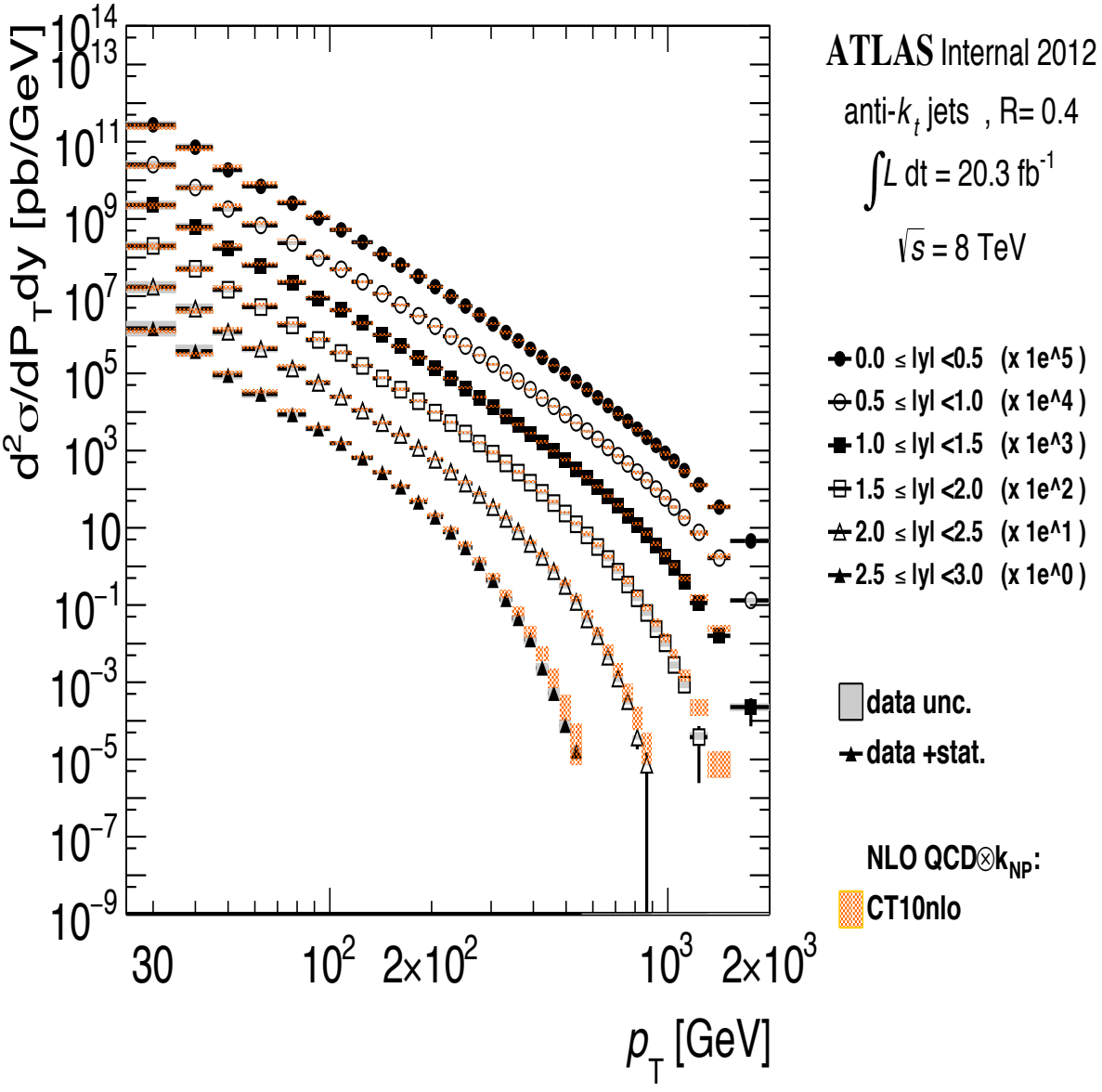


Figure 4.3: Inclusive double-differential cross sections for anti-k_t jets with radius parameter R = 0.4, shown as a function of p_T and $|y|$. To help visibility, the cross sections are multiplied by the factors indicated in the legend. The black error bars indicates the statistical uncertainty of data. The gray-shaded band indicates the sum in quadrature of the experimental systematic uncertainties of data. For comparison, the NLO QCD predictions of NLOJet++ using the CT10 PDF set, corrected for non-perturbative, are included. The hatched orange band shows the uncertainty associated to the theory predictions.

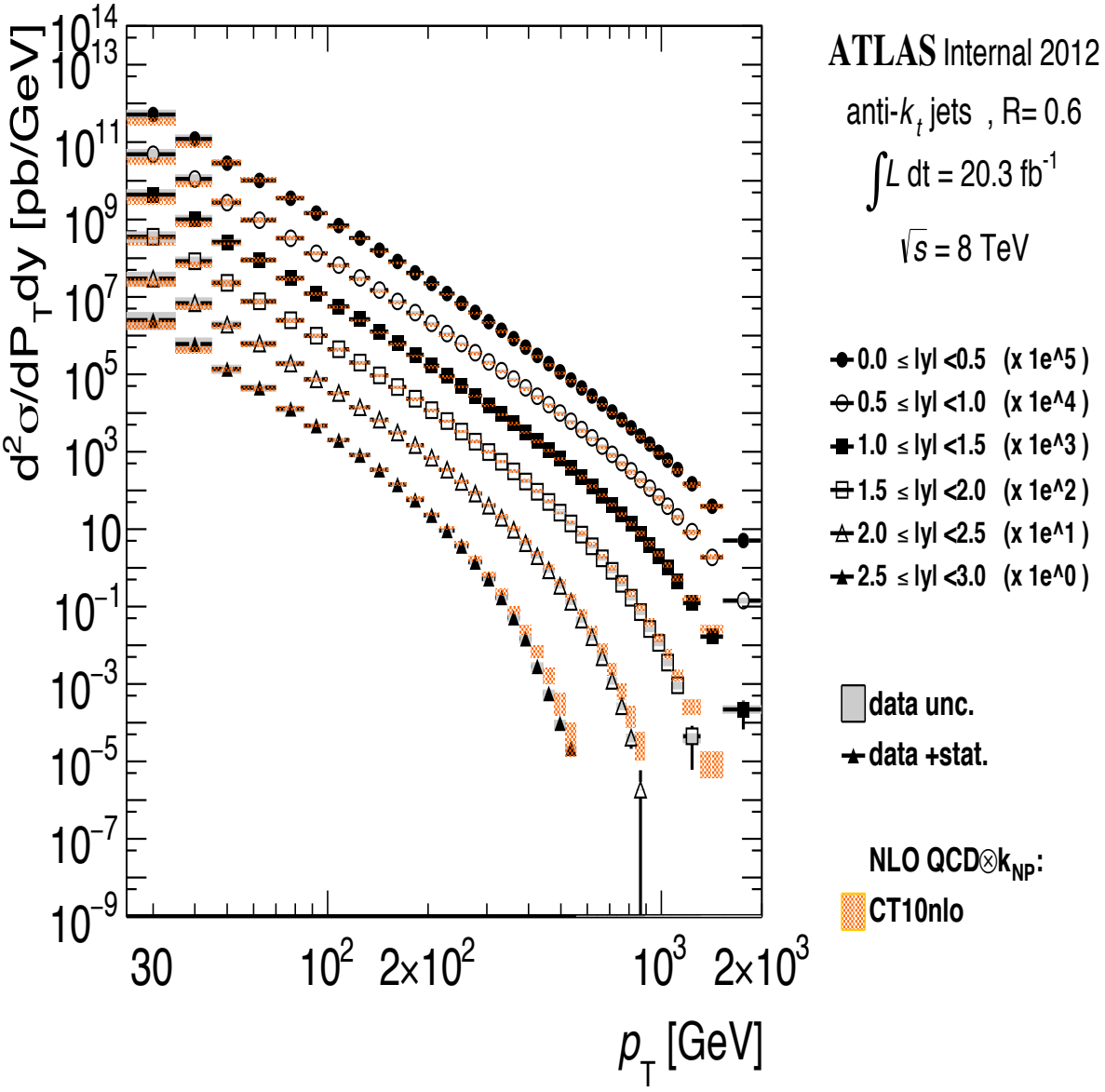


Figure 4.4: Inclusive double-differential cross sections for anti-k_t jets with radius parameter R = 0.6, shown as a function of p_T and $|y|$. To help visibility, the cross sections are multiplied by the factors indicated in the legend. The black error bars indicates the statistical uncertainty of data. The grey-shaded band indicates the sum in quadrature of the experimental systematic uncertainties of data. For comparison, the NLO QCD predictions of NLOJet++ using the CT10 PDF set, corrected for non-perturbative, are included. The hatched orange band shows the uncertainty associated to the theory predictions.

4.14, theory predictions from the various PDF sets are compared to data in all transverse momentum ranges, combining events from the no-pileup, low-μ and high-μ runs. The relative total uncertainties for the various PDF sets is quite similar, as already seen in figure 3.15 in section 3.11. The dashed azure band represents ratio between theory and data(no pile-up) which was not added in the ATLAS publication. To remind a large luminosity uncertainty, it shows with orange band. The systematic uncertainty for R = 0.6 is in general smaller than the corresponding one for R=0.4, and the agreement with data is usually better. Comparing the various PDF sets it can be observed that, while predictions are similar in the low-p_T region, CT10 does predict a higher cross-section than data at high p_T.

4.2 Conclusions for the inclusive jet cross-section

The inclusive jet cross-section in proton-proton collisions at $\sqrt{s} = 8$ TeV has been measured for jets reconstructed with the anti-k_t algorithm with jet radius parameter values of $R = 0.4$ and $R = 0.6$ in the kinematic region of the jet transverse momentum $p_T \geq 25$ GeV and jet rapidities $|y| < 3$. The measurement is based on the data collected with the ATLAS detector during LHC operation in 2012, corresponding to a total integrated luminosity of 193.3 nb^{-1} for lower μ runs and 20.3 fb^{-1} for high μ run. The cross-sections are measured double differentially in the jet transverse momentum and rapidity. The measurement is extended to 25 GeV in jet transverse momentum using the no pile-up run. The dominant systematic uncertainty arises from the jet energy calibration. Compared to previous jet cross-section measurements at $\sqrt{s} = 7$, TeV a significant reduction of the cross-section uncertainties is achieved, also thanks to the use of the low-μ run for the low transverse momentum region, since the high mu run sample was used between 25-100 GeV in the measurement at $\sqrt{s} = 7$ TeV [85].

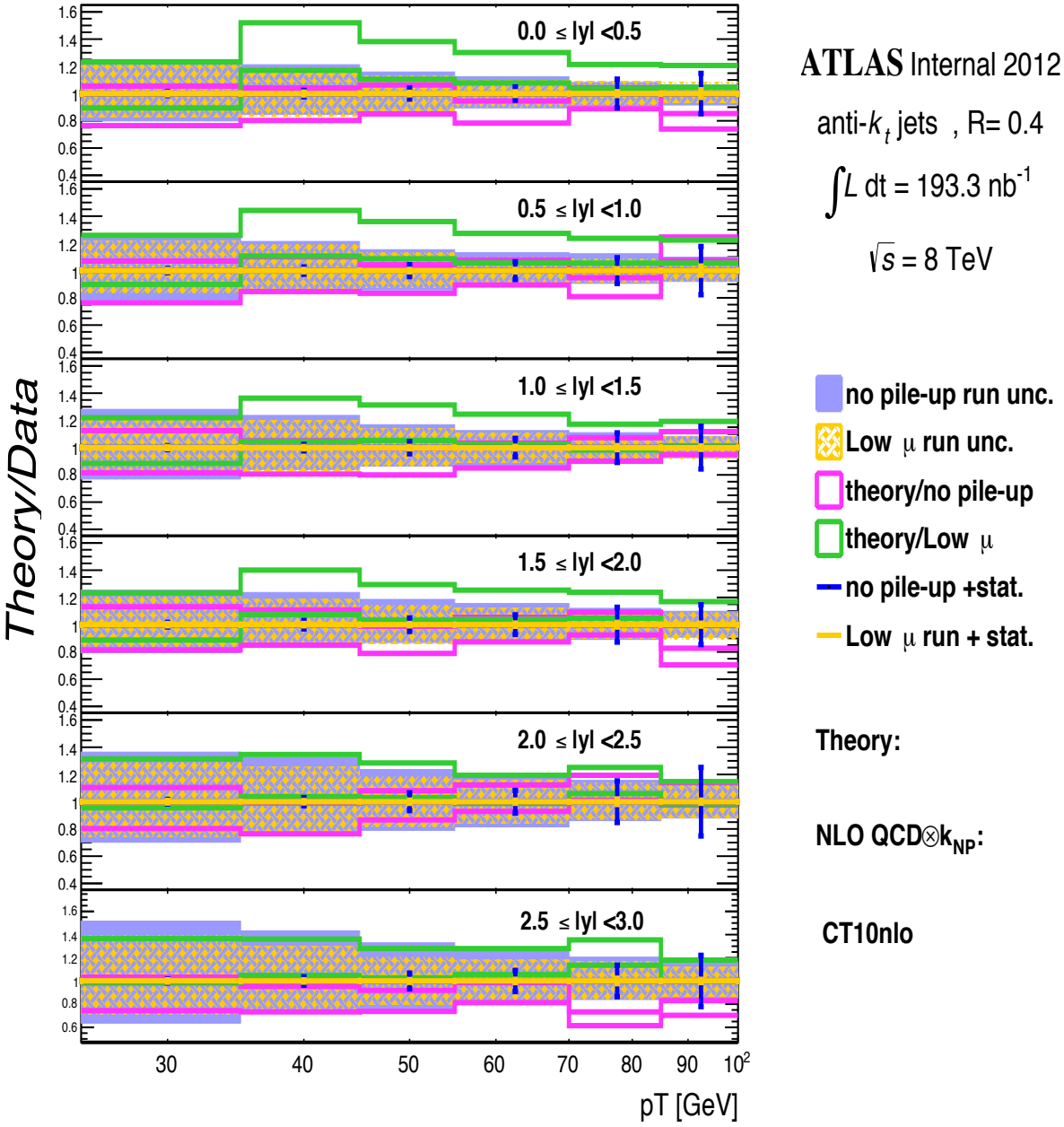


Figure 4.5: Ratio of the NLO QCD predictions calculated for CT10nlo PDF-set multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.4, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as error bars, blue line for no pile-up and orange line for low μ run. Total systematic uncertainties are shown with violet band and dashed orange band for respectively no pile-up and low μ run. The pink lines are shown the ratio between theory and data(no pile-up), green lines are shown ratio between theory(CT10) and data(low μ run)

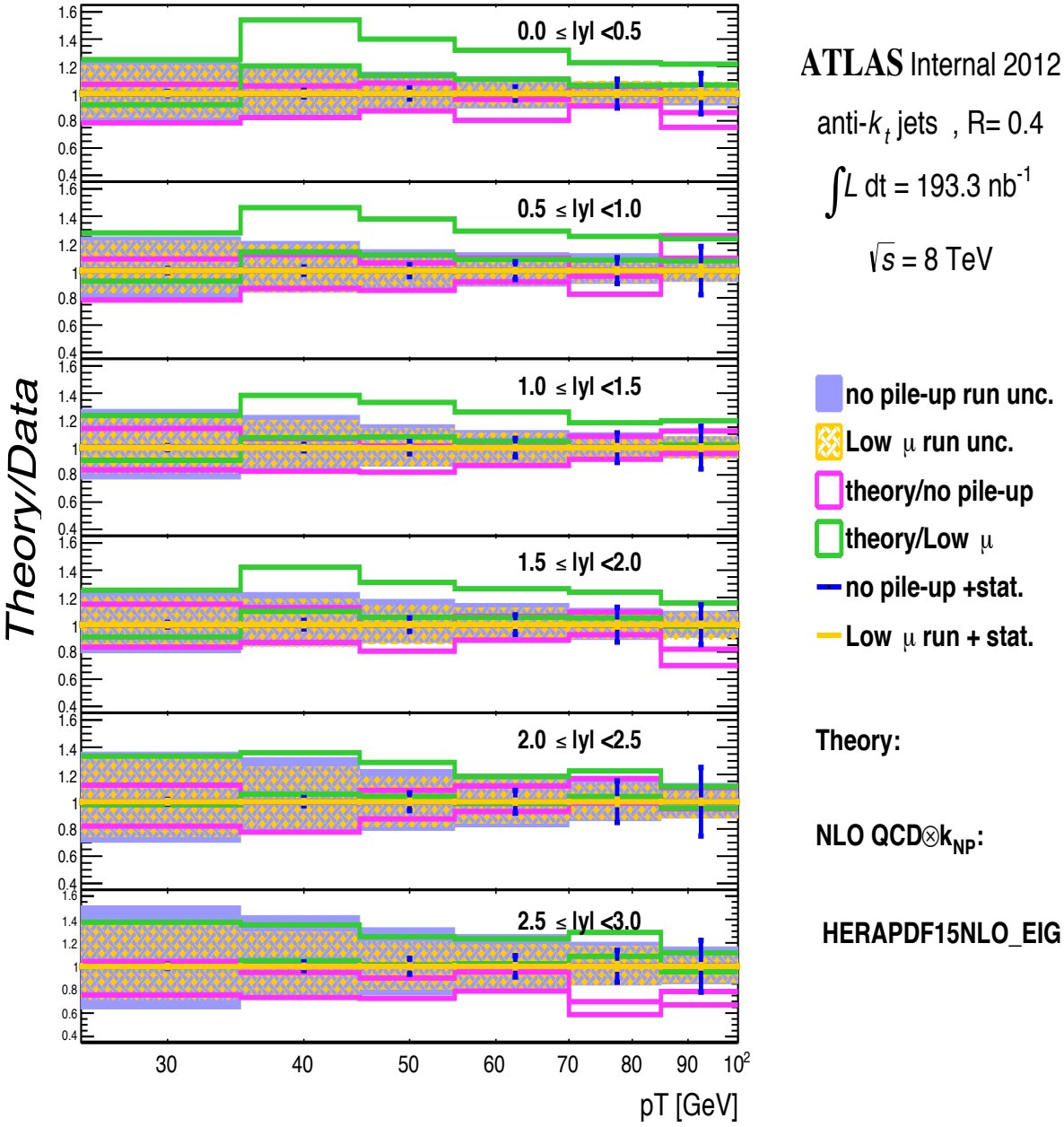


Figure 4.6: Ratio of the NLO QCD predictions calculated for HERAPDF15NLO_EIG PDF-set multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.4, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as error bars, blue line for no pile-up and orange line for low μ run. Total systematic uncertainties are shown with violet band and dashed orange band for respectively no pile-up and low μ run. The pink lines are shown the ratio between theory and data(no pile-up), green lines are shown ratio between theory(HERAPDF15NLO_EIG) and data(low μ run)

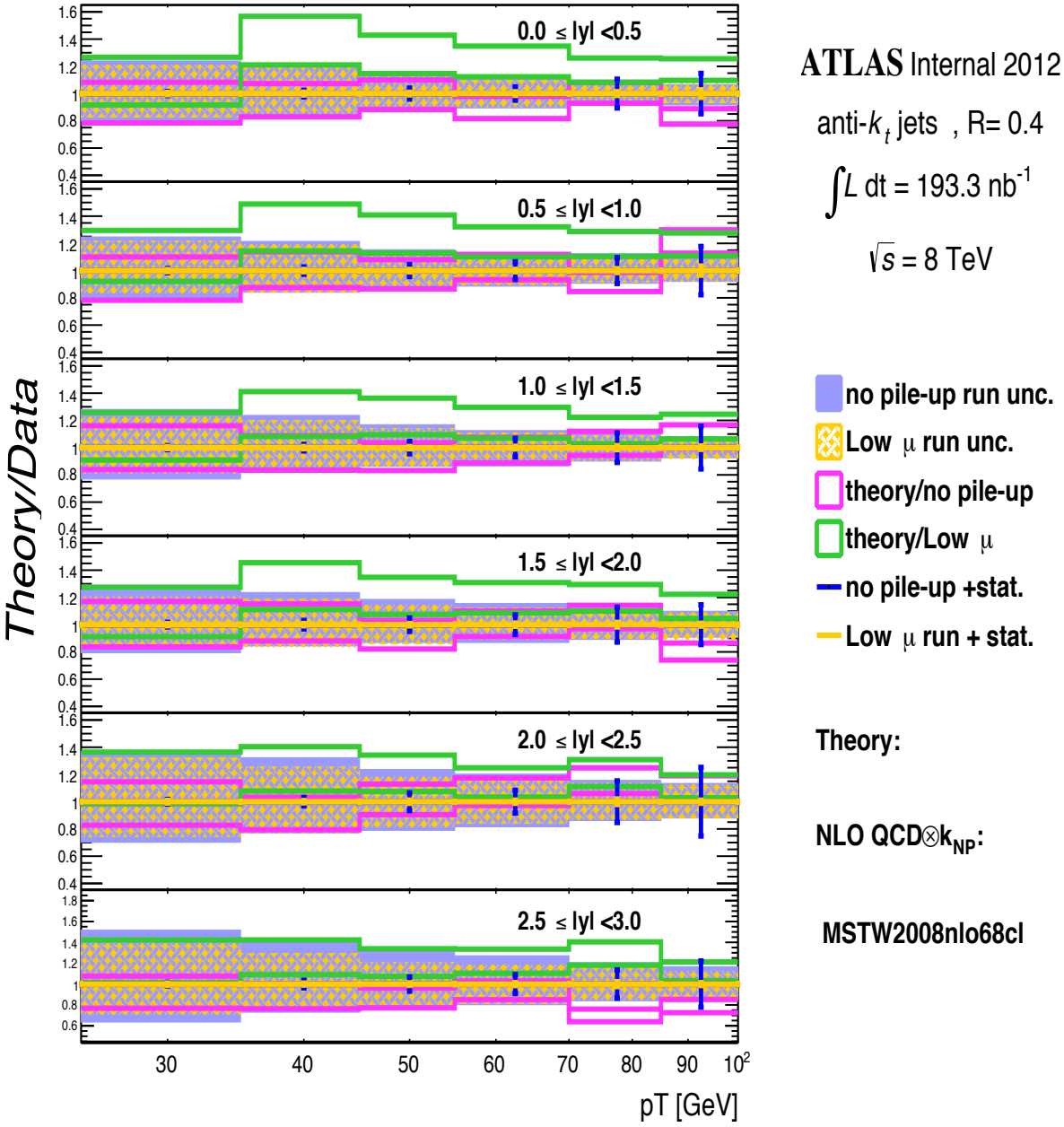


Figure 4.7: Ratio of the NLO QCD predictions calculated for MSTW2008nlo68cl PDF-set multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.4, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as error bars, blue line for no pile-up and orange line for low μ run. Total systematic uncertainties are shown with violet band and dashed orange band for respectively no pile-up and low μ run. The pink lines are shown the ratio between theory and data(no pile-up), green lines are shown ratio between theory(MSTW2008nlo68cl) and data(low μ run)

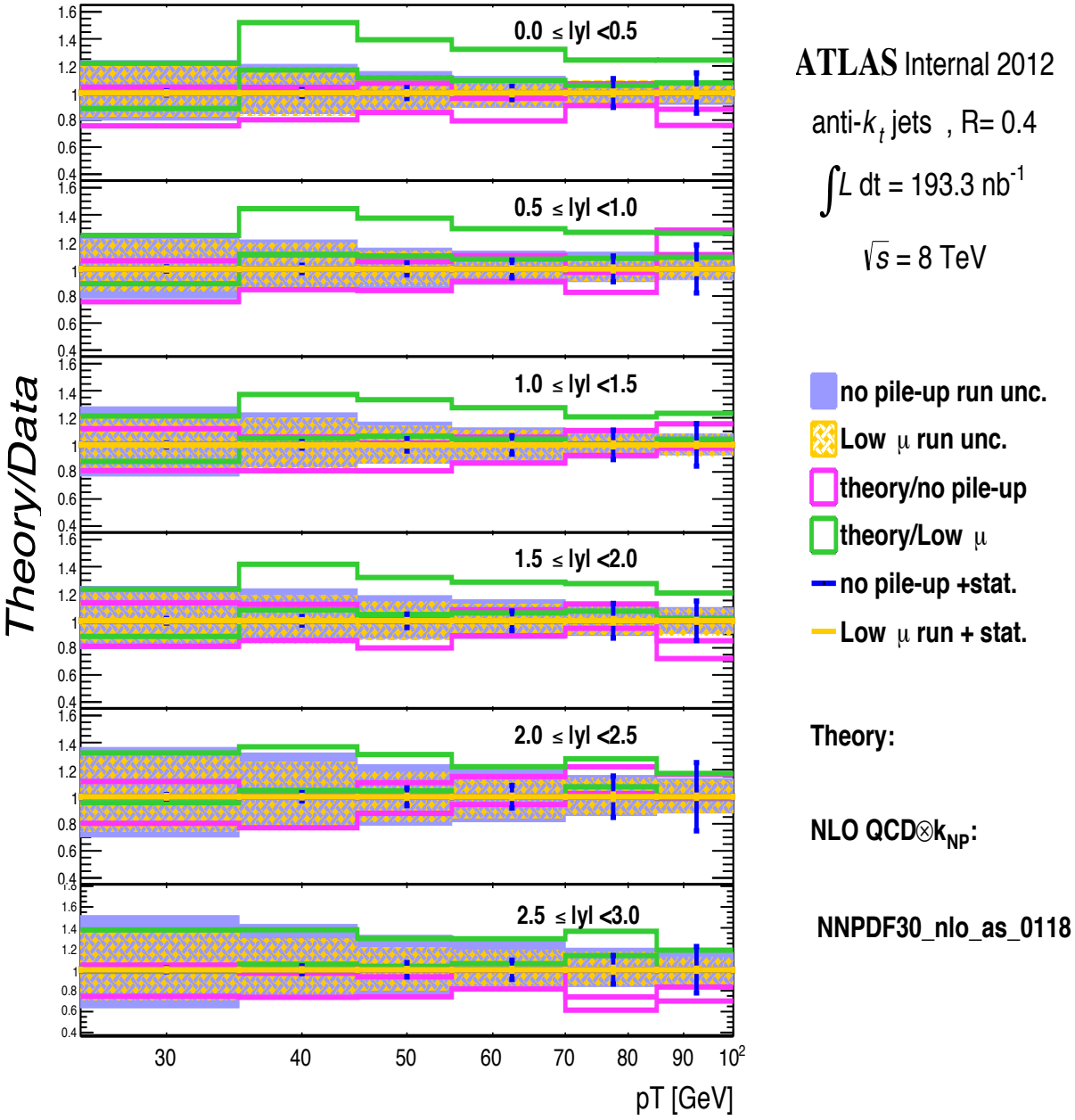


Figure 4.8: Ratio of the NLO QCD predictions calculated for NNPDF30_nlo_as_0118 PDF-set multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.4, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as error bars, blue line for no pile-up and orange line for low μ run. Total systematic uncertainties are shown with violet band and dashed orange band for respectively no pile-up and low μ run. The pink lines are shown the ratio between theory and data(no pile-up), green lines are shown ratio between theory(NNPDF30_nlo_as_0118) and data(low μ run)

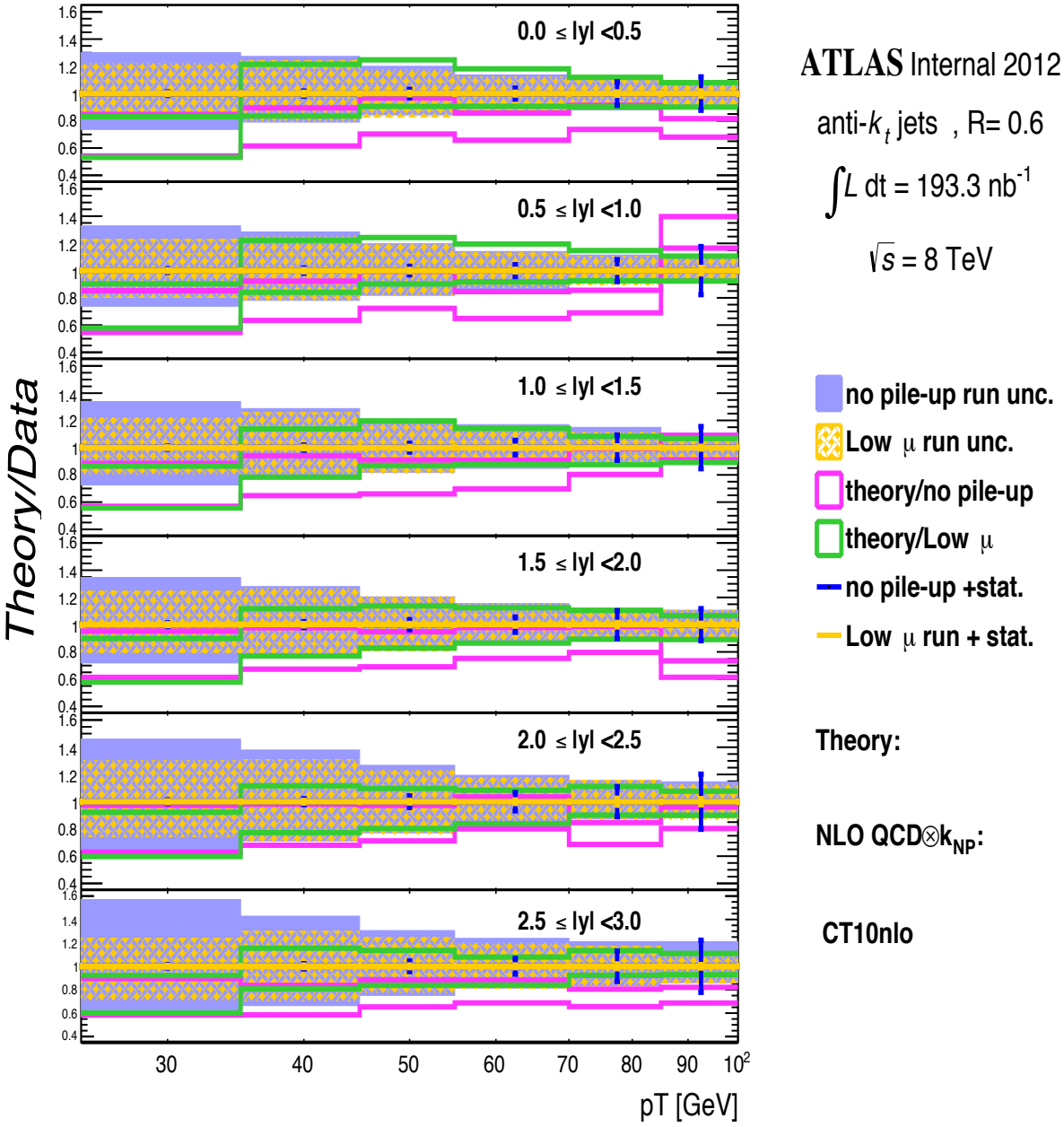


Figure 4.9: Ratio of the NLO QCD predictions calculated for CT10nlo PDF-set multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.6, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as error bars, blue line for no pile-up and orange line for low μ run. Total systematic uncertainties are shown with violet band and dashed orange band for respectively no pile-up and low μ run. The pink lines are shown the ratio between theory and data(no pile-up), green lines are shown ratio between theory(CT10) and data(low μ run)

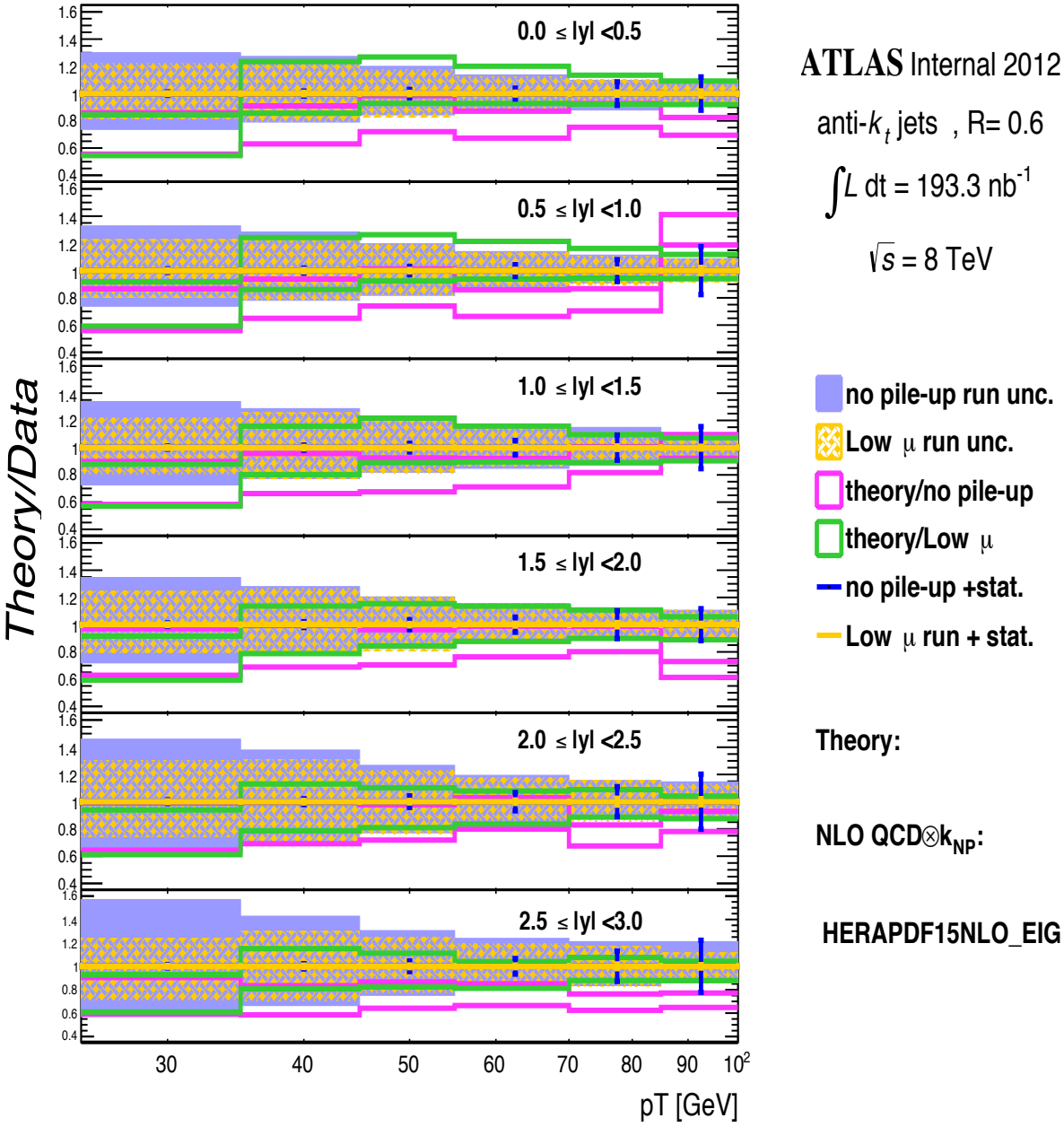


Figure 4.10: Ratio of the NLO QCD predictions calculated for HERAPDF15NLO_EIG PDF-set multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.6, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as error bars, blue line for no pile-up and orange line for low μ run. Total systematic uncertainties are shown with violet band and dashed orange band for respectively no pile-up and low μ run. The pink lines are shown the ratio between theory and data(no pile-up), green lines are shown ratio between theory(HERAPDF15NLO_EIG) and data(low μ run)

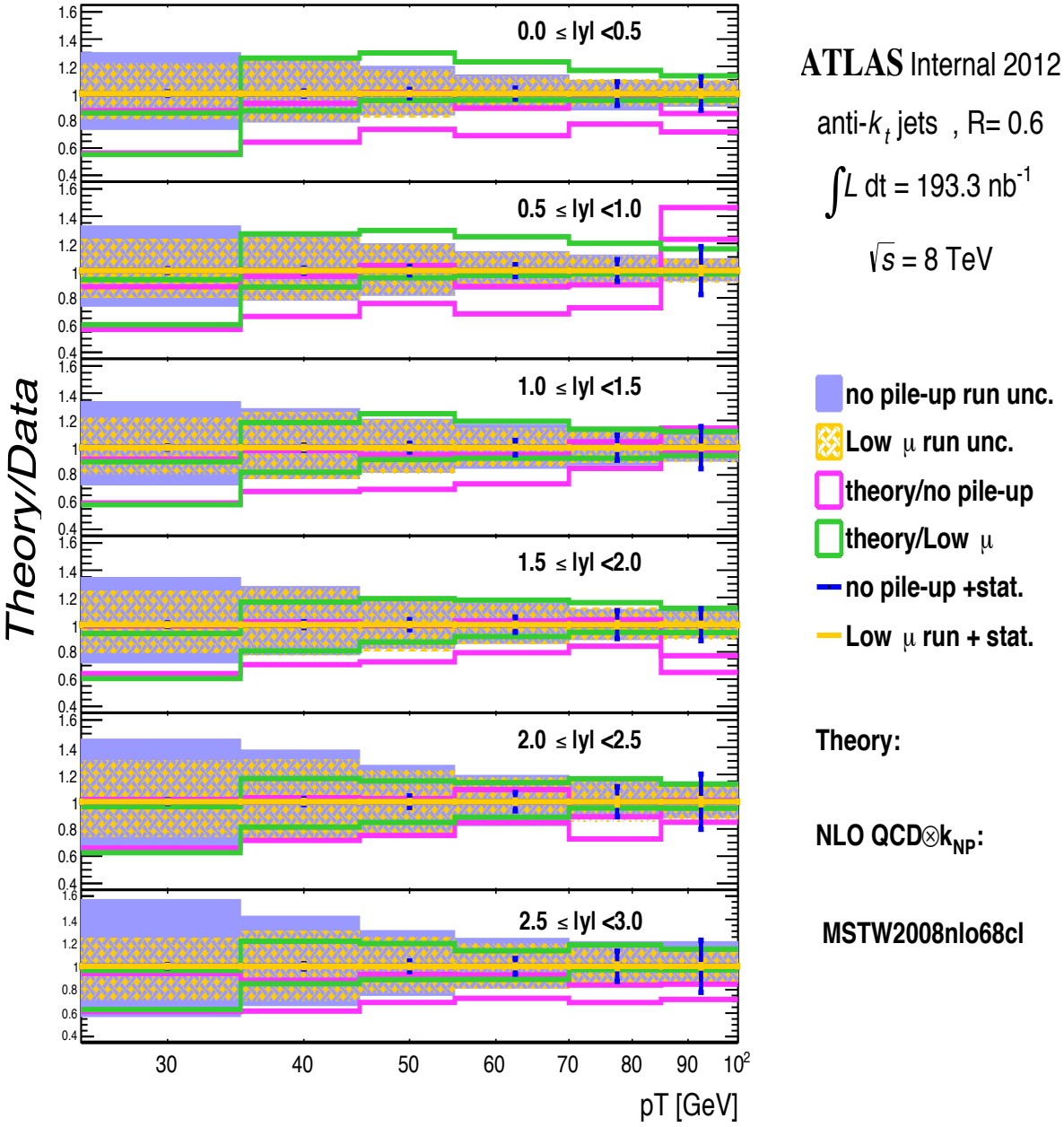


Figure 4.11: Ratio of the NLO QCD predictions calculated for MSTW2008nlo68cl PDF-set multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.6, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as error bars, blue line for no pile-up and orange line for low μ run. Total systematic uncertainties are shown with violet band and dashed orange band for respectively no pile-up and low μ run. The pink lines are shown the ratio between theory and data(no pile-up), green lines are shown ratio between theory(MSTW2008nlo68cl) and data(low μ run)

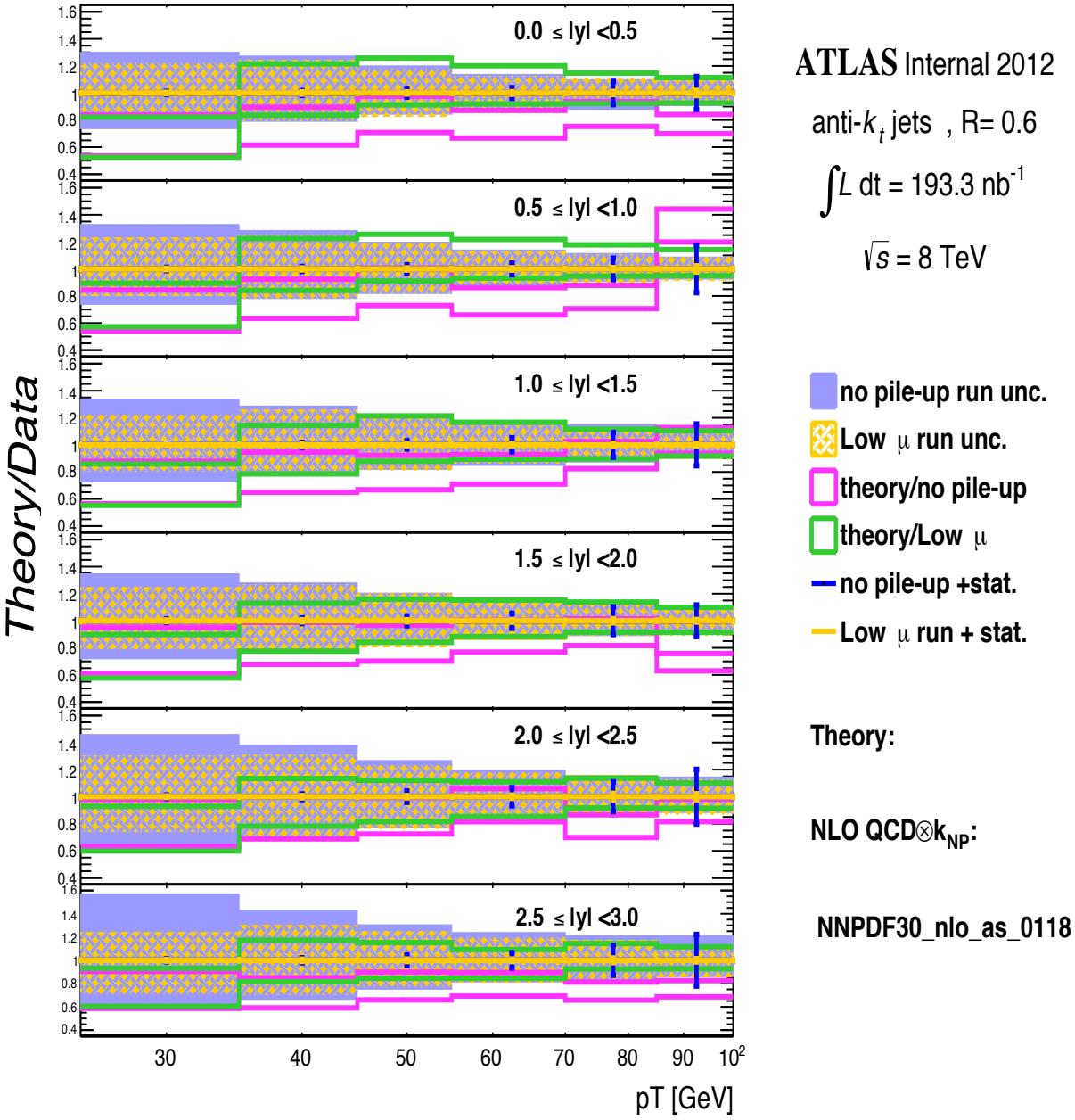


Figure 4.12: Ratio of the NLO QCD predictions calculated for NNPDF30_nlo_as_0118 PDF-set multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.6, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as error bars, blue line for no pile-up and orange line for low μ run. Total systematic uncertainties are shown with violet band and dashed orange band for respectively no pile-up and low μ run. The pink lines are shown the ratio between theory and data(no pile-up), green lines are shown ratio between theory(NNPDF30_nlo_as_0118) and data(low μ run)

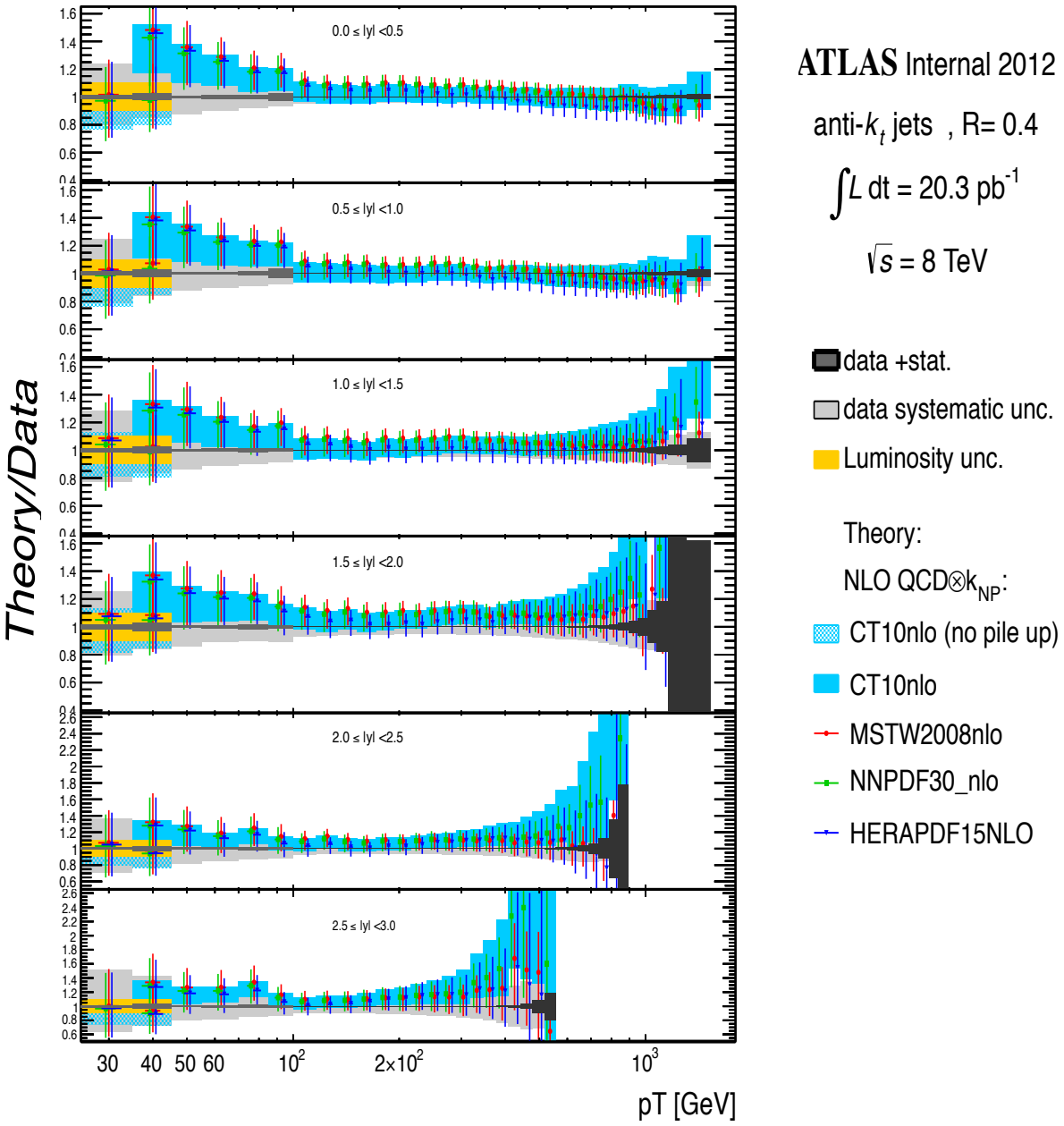


Figure 4.13: Ratio of the NLO QCD predictions calculated with various PDF-sets, (CT10, HERAPDF1.5, MSTW 2008, NNPDF 3.0) multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.4, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as dark grey band. Total data systematic uncertainties are shown with grey band. Orange band shows the luminosity uncertainty, it is just illustrated at one bin between 35-45 GeV in order to see differences between no pile-up and low μ run. Ratio of theory to inclusive-jet cross section measured in data is illustrated with azure band. Ratio of theory to no pile-up run in data is dashed azure band. Due to see much clearer, x error of theory has been made smaller and while NNPDF 3.0 moved to left, HERAPDF moved to right by 2% shift away from central.

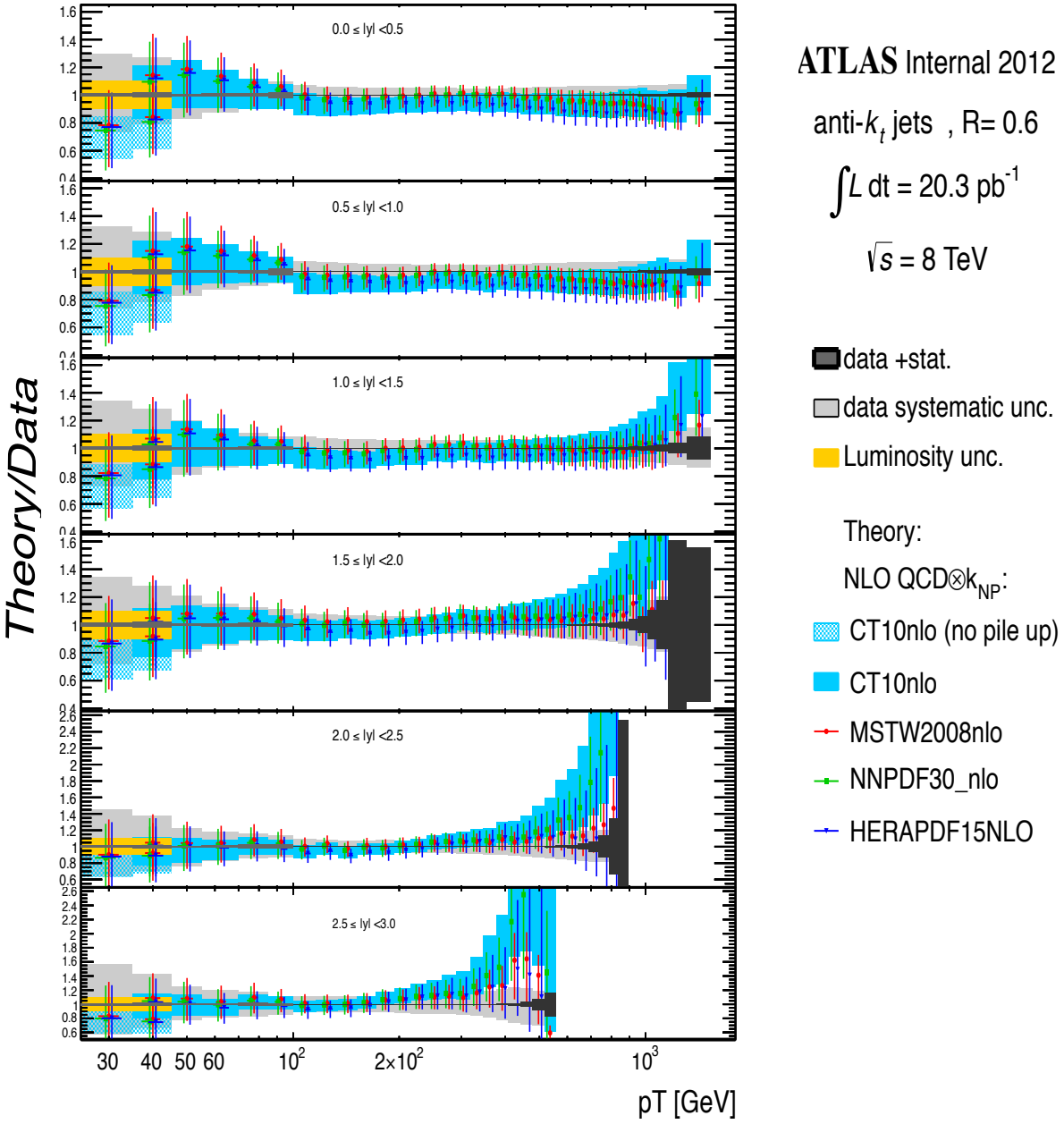


Figure 4.14: Ratio of the NLO QCD predictions calculated with various PDF-sets, (CT10, HERAPDF1.5, MSTW 2008, NNPDF 3.0) multiplied with the non-perturbative correction (k_{NP}) to the inclusive-jet cross section measured in data for anti-k_t jets with radius parameter R = 0.6, in different ranges of $|y|$. The relative statistical uncertainties of the measurements are indicated as dark grey band. Total data systematic uncertainties are shown with grey band. Orange band shows the luminosity uncertainty, it is just illustrated at one bin between 35-45 GeV in order to see differences between no pile-up and low μ run. Ratio of theory to inclusive-jet cross section measured in data is illustrated with azure band. Ratio of theory to no pile-up run in data is dashed azure band. Due to see much clearer, x error of theory has been made smaller and while NNPDF 3.0 moved to left, HERAPDF moved to right by 2% shift away from central.

This measurement has a finer binning in p_{T}, thus giving more precise information on the p_{T}-dependence of the inclusive jet cross-section. Fixed-order NLO perturbative QCD calculations have been corrected for both non-perturbative effects, but electroweak effects have shown to be negligible at low transverse momentum, so have only been considered for the high-μ part of the data. Several NLO PDF sets are used in the theory predictions for this comparisons. Most of the NLO pQCD predictions are in good agreement with the measurement, confirming that perturbative QCD can describe jet production; some tension is however present in the lowest part of the spectrum for jets reconstructed with R=0.6, possibly indicating a not perfect understanding of the soft effects for larger jet cones.

Chapter 5

Study of Quark-Gluon Jet Discrimination

5.1 Introduction

Billions of jets a second are produced at the Large Hadron Collider, and the majority of these jets will be originated by the showering of a parton, either a quark or a gluon. The possibility of distinguishing between jets originated by quarks and those originated by gluons would be beneficial to several analyses, where usually the signal is composed of quarks and the background of gluons (examples of these cases are R-parity violating SUSY models, leptophobic $Z^{'}$ or $W^{'}$, jets from Higgs VBF; while in other cases like gluino-pair production, gluons are a signal). Signal and background can be discriminated using observables like jet mass, which are strongly correlated with flavor. Moreover, separating quark-initiated from gluon-initiated jets can also dramatically improve the reach of searches for new physics models where no clear resonance is present, such as contact interactions at the LHC. When the scale Λ is much higher than the measured invariant mass or p_T, no peak observation is expected, and the only consequence of these interactions would be a rise of the cross-section for high invariant masses. However, the high invariant mass region, that also corresponds to high-x in the proton PDF's, is also poorly constrained, and in case an excess is observed with respect to the predictions based

on current PDF's, it is not obvious whether it should be seen as an indication of new physics or just of our poor knowledge of PDF's, and included in the next fit. The possibility of distinguishing quark-initiated from gluon-initiated jets would add further constraints to the dijet system, therefore improving the ability to distinguish new physics effects from just a harder quark spectrum inside the proton.

Despite the obvious advantages of a tool able to discriminate between quark and gluon jets, producing such an algorithm is not straightforward since the discrimination between quark-initiated and gluon-initiated jets relies on soft-physics properties of the jet, that are not usually well reproduced by the Monte Carlo codes. The high luminosity collected in by the LHC allows the selection of large samples of relatively pure quark- or gluon-initiated jets, so extracting these distribution from data decreases the dependence from the Monte Carlo modelling, overcoming at least partially this limitation. Processes leading to relatively pure samples of light quarks or gluons have been highlighted in several phenomenological works, like [86] [87] [88] [89]. For instance, it was found that at the 7TeV LHC, jets with transverse momentum above 200 GeV coming from the $pp \to \gamma$+2jets sample are quark-initiated in 98% of the cases, for a cross section of 5 pb; on the other hand, the third jet in a $pp \to$ 3jets process can provide a 90% gluon purity above 200 GeV [90].

For the analysis of the 2011 dataset, a tool has been built using likelihood functions, based on two different variables: the number of tracks in a jet and the jet width calculated from tracks. In this tool, a 50% quark jet efficiency can be achieved with about 80%-90% of gluon jets suppression [91]. Using this tool, jets can be labelled as quark or gluon and then dijet events labelled as quark-quark, quark-gluon or gluon-gluon. Separate detector-level mass spectra for the three flavor combinations can be plotted in bins of y^*, and compared to Standard Model QCD and with contact-interaction models using different Λ scales. In the first part of this chapter, track variables for quark-gluon separation tool are described, and then a comparison will be shown between

dijet invariant masses in data and Monte Carlo for models with and without contact interactions, using the quark-gluon separation tool.

The 2012 analysis, described in the second part of this chapter, extends this previous study to the inclusion of calorimeter-based variables on top of the track-based ones used in 2011. These variables are studied in bins of p_T and η because the quark and gluon fractions, as well as jet properties, can be strongly dependent on kinematics. The aim of this study is to provide good templates for jet-shape distributions using W+jet (where the W boson decays into a muon and a neutrino) and dijet events. In addition, also γ+jet events are used to improve separation power between the two jet categories.

5.2 Event and Jet Selection for the Quark-Gluon separation tool in 2011

The 2011 Quark-Gluon tagger has been applied on data collected at a centre-of-mass energy of 7 TeV, and compared to dijet Monte Carlo events generated with the Pythia 6 generator with AMBT1 tune and MRST2007lomod PDF [92]. The luminosity collected during this year has been 4.7 $\pm$0.2 fb^{-1} using periods D-M. Since during data taking some instability of the LAr calorimeter was observed, events containing a LAr noise burst, or data corruption at the level $larError != 2$ are rejected, and as well as those containing $tileError != 2$ (see more explanation in section 3.4)The kinematic region considered in this analysis for jets is $p_T > 60$ GeV and $|\eta| < 2.1$, to be in the full acceptance of the tracking detector. Only jets falling in the cleaning definition of 'GOOD' (so, not bad neither ugly), are accepted. Additionally, the two leading jet in the event must pass the "Medium" cleaning quality criteria, to reject cosmic rays, beam halo, and detector noise. Selected jets are calibrated locally applying a jet energy scale (JES) correction, that also accounts for bad channels.

The vertex selection requires at least three primary tracks with $p_T > 0.5$, to remove events coming from cosmic-ray muons and other non-collision back-

grounds. Tracks with $p_T > 1$GeV and $|\eta| < 2.4$ are associated to jets using a ΔR matching method which provides to find the closest tracks to jets. At lower p_T, tracks are more sensitive to fragmentation and underlying events. Tracks are selected to suppress pile-up effects with the cuts in table 5.1.

$Z_0 sin(e) < 1.0$mm	nPixHits ≥ 1
$p_T > 1$GeV	nSCTHits ≥ 6
$D_0 < 1.0$mm	chi2/ndof ≤ 3.0

Table 5.1: The cuts for tracks.

5.2.1 Truth Jets and Flavour Labelling

In order to assign a flavour to a reconstructed jets, truth jets reconstructed by clustering stable hadrons in the MC are used. To define a jet as light quark, c-quark or b-quark or gluon-initiated, the event record is inspected, and the highest energy partons within the jet cone, ΔR, is used to label that jet.

5.2.2 Track-based variables for quark-gluon separation

In the 2011 analysis, track information is used for its robustness with respect to pileup, due to the possibility of assigning tracks to the primary vertex. The two variables used in the analysis are:

- Number of tracks in the jet:

$$n_{trk} = \sum_{\text{trk}\in\text{jet}} \tag{5.1}$$

- Track-based p_T weighted width of the jet:

$$w_{trk} = \frac{\sum_{\text{trk}\in\text{jet}} p_{\text{T,trk}} \Delta R_{\text{trk,jet}}}{\sum_{\text{trk}\in\text{jet}} p_{\text{T,trk}}} \tag{5.2}$$

Figure 5.1 and 5.2 show respectively track width and track size for jets selected with different p_T cut in the central $|\eta|$ region. It is expected that

discrimination between quark and gluon jets decreases by increasing p_T as shown by these figures. At high pT , quark jets behave like gluon ones because they collimate with the direction of the original parton. The number of tracks gives a better separation power compared with track width. The two variables are then combined in a likelihood, and jets are identified as originating from quarks or gluons after a cut on this likelihood that has a 50% quark efficiency and 80% gluon rejection.

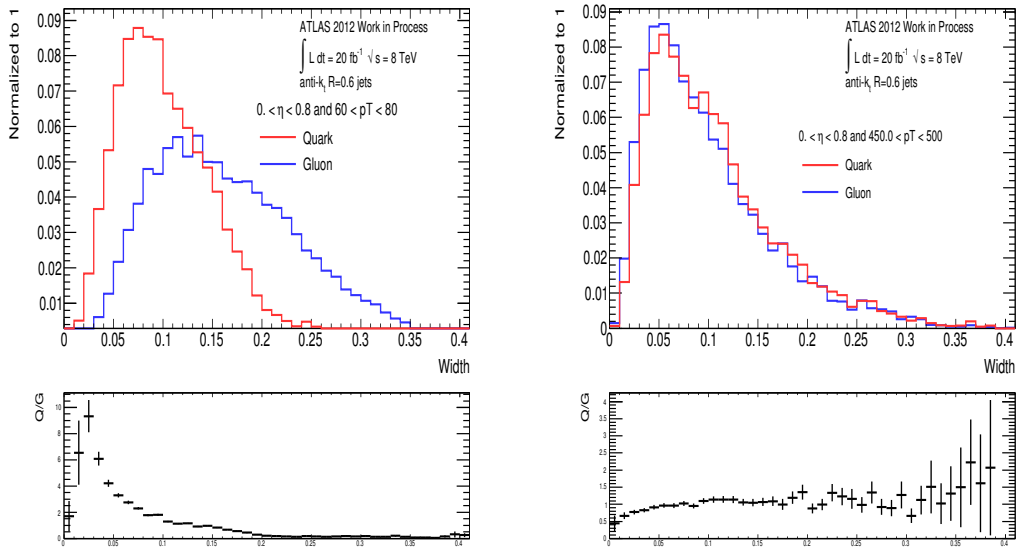


Figure 5.1: The track width at different p_T range at $|\eta| < 0.8$, events at detector level are taken from MC generator. Left figure is the width at $60 < p_T < 80$, right one is at $450 < p_T < 500$.

5.2.3 Selection of dijet events

The dijet double-differential cross section based on 2011 data sample is measured as a function of the dijet invariant mass, m_{12} in bins of half the rapidity separation of the two leading jets y^*, $|y_1 - y_2|$. Data is compared to Monte Carlo simulation for SM QCD, and different values of the compositeness parameter Λ: 4TeV, 6TeV and 8TeV. Also the variables used for quark-gluon separation, track width and number of tracks are compared to SM QCD, and simulations for compositeness with Λ=4TeV and 6TeV.

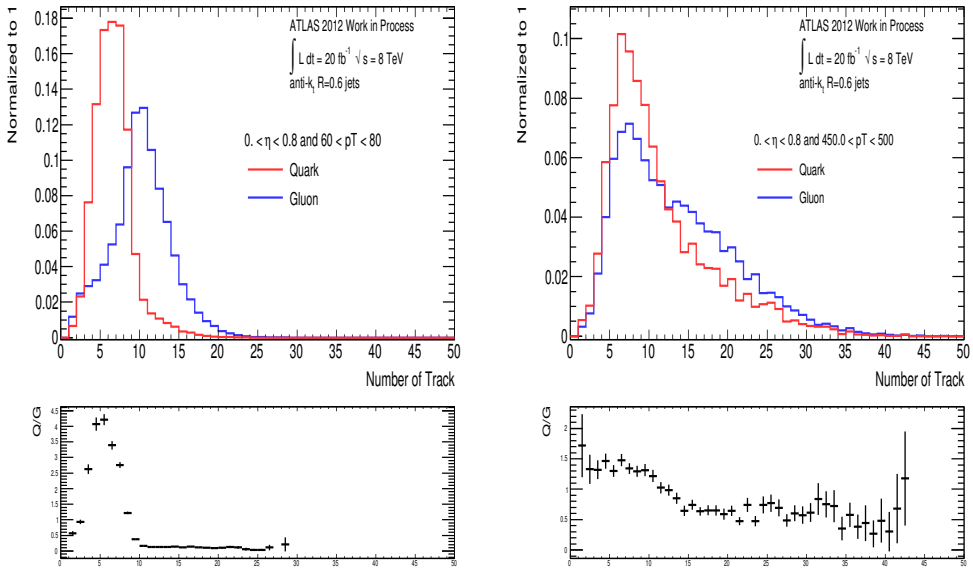


Figure 5.2: The number of track at different p_T range at $|\eta| < 0.8$, events at detector level are taken from MC generator. Left figure is the width at 60 $< p_T <$80, right one is at $450 < p_T < 500$.

For the dijet selection, jets are ignored if one of the two leading jets falls in region of $-0.88 < \phi < -0.5$ and rapidity compatible with the LAr hole. The first leading jet is required to have $p_T > 100$ GeV, while the second leading jet must have $p_T > 50$ GeV. These two jets are asymmetric to improve stability of NLO calculation The azimuthal angle between the two leading jets must be $\Delta\phi > 2.5$ [93].

This data is only selected using central jet triggers at table 5.2, corresponding to a series of p_T bins where the various triggers are more than 99% efficient in the rapidity region $|y| < 3.0$. This choice removes the need to correct for trigger efficiency, and allows the use of an equivalent luminosity for each transverse momentum interval, calculated from the prescale of the corresponding trigger [85].

[1] **EF_jX_a4tc_EFFS:**, EFFS marks jet and combined jet + E_T trigger chains which use the full-scan algorithm at EF at p_T ¿X. a4tc means that jets are reconstructed using the anti-kt algorithm with topological cell energy clusters as signal input with a radius of R=0.4.

Leading Jet p_T	**Trigger**
40 GeV	EF_j10_a4tc_EFFS[1]
60 GeV	EF_j20_a4tc_EFFS
80 GeV	EF_j30_a4tc_EFFS
110 GeV	EF_j55_a4tc_EFFS
160 GeV	EF_j75_a4tc_EFFS
210 GeV	EF_j100_a4tc_EFFS
260 GeV	EF_j135_a4tc_EFFS

Table 5.2: Leading jet p_T bins used to select the trigger used for the di-jet samples.

5.2.4 Comparison of SM QCD with Contact Interactions

Figure 5.3 shows the transverse momentum distribution of all jets (left) and of the leading jet (right) in the event in the case of QCD and contact interactions with the Λ parameter set at 4 and 6 TeV. The presence of contact interaction produces a gentle change in the slope and increases the rate relative to QCD at high p_T. As already mentioned in the introduction, it is possible that the presence of contact interactions is mistaken as harder Pdfs for large x values, and the existence of new physics may be difficult to prove.

The different mix of quarks and gluons present in new physics scenarios may allow an improved separation when the quark-gluon tagger is used. Figure 5.4 shows the distribution of the number of tracks for jets labelled in the Monte Carlo as originated by quarks on the left and by gluons on the right. Figure 5.5 shows instead the distribution of the track width, also for quark-labelled and gluon-labelled jets.

Figures in 5.6 show the dijet invariant mass distribution divided into gluon-gluon(M_{gg}), quark-quark(M_{qq}) and quark-gluon(M_{qg}) tagged events, after a cut on the likelihood is performed. Events are shown in different dijet y^* intervals, and for the full 2011 data sample. Higher values of Λ produce mass spectra closer to those predicted by SM QCD: we expect that the high-mass

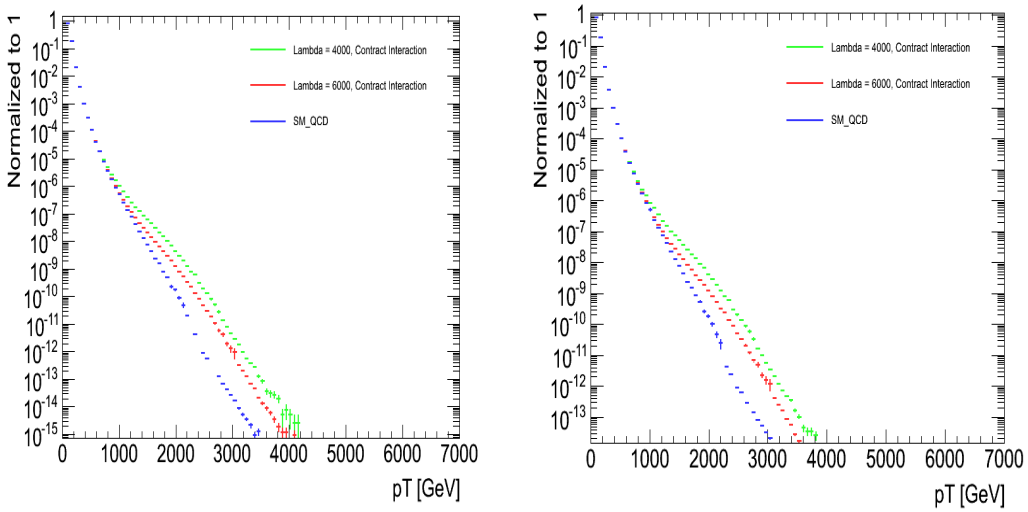


Figure 5.3: The left figure shows the distribution of the transverse momentum of all jets passing the 2011 dijet selection, for SM QCD Monte Carlo (blue line), and for Contact Interactions with Lambda = 4 TeV (green line) and Lambda = 6 TeV (red line). The right plot shows the distribution of the transverse momentum of leading jets only. For a proper comparison, the distributions are normalised to unity.

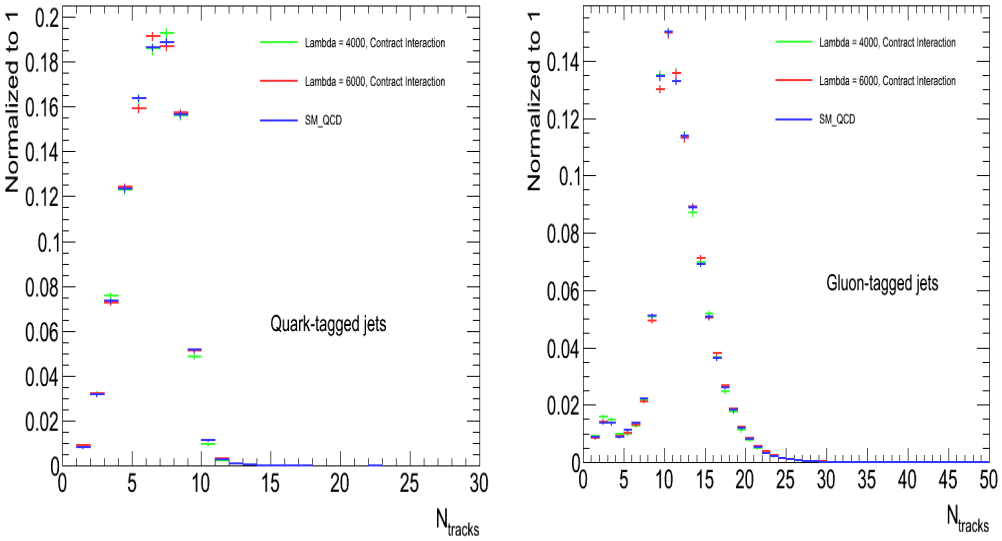


Figure 5.4: The number of the tracks inside the quark tagged jets (left) and gluon tagged jets (right) for SM QCD (blue line), and Contact Interactions with Lambda = 4 TeV (green line) and Lambda = 6 TeV (red line). They are normalised to one for a reasonable comparison.

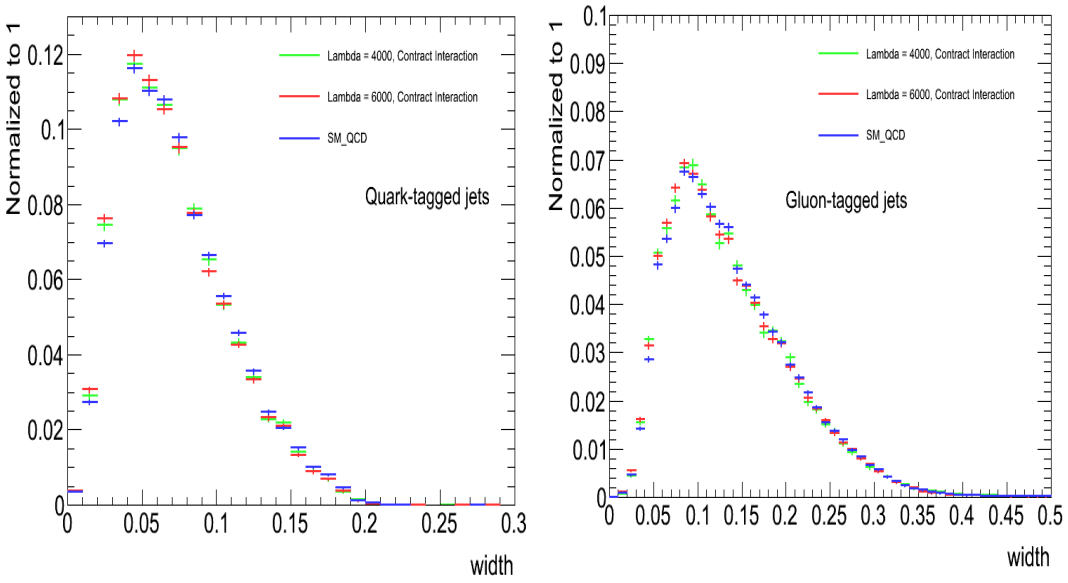


Figure 5.5: The track width inside the quark tagged jets (left) and gluon tagged jets (right) at $\sqrt{s} = 7$TeV in 2011 for SM QCD (blue line), and Contact Interactions with Lambda = 4 TeV (green line) and Lambda = 6 TeV (red line)They are normalised to one for a reasonable comparison.

cross section for Λ_{4000} is larger than that for Λ_{6000}, which is larger than that for Λ_{8000}. However this expectation does not occur clearly on M_{qq} and M_{qg}. The reason of this case is that quark and gluon jets were not well separated and they may be mistagged by tool. The previous exclusion limits on the compositeness scale for quark CI ranges from 2.8 to 3.1 TeV at 95% CL [94] [94], so this dataset has the power of setting more stringent limits because the increases on contact interaction cross-section is not large. Therefore it was chosen not to use it to set the final limit on the ATLAS published result. The tool for 2011 should be improved, and the higher statistics present in the 2012 dataset allows the use of better and smoother purified samples. Therefore, my thesis work has continued with the study of Quark-Gluon Jet discrimination using the 2012 dataset.

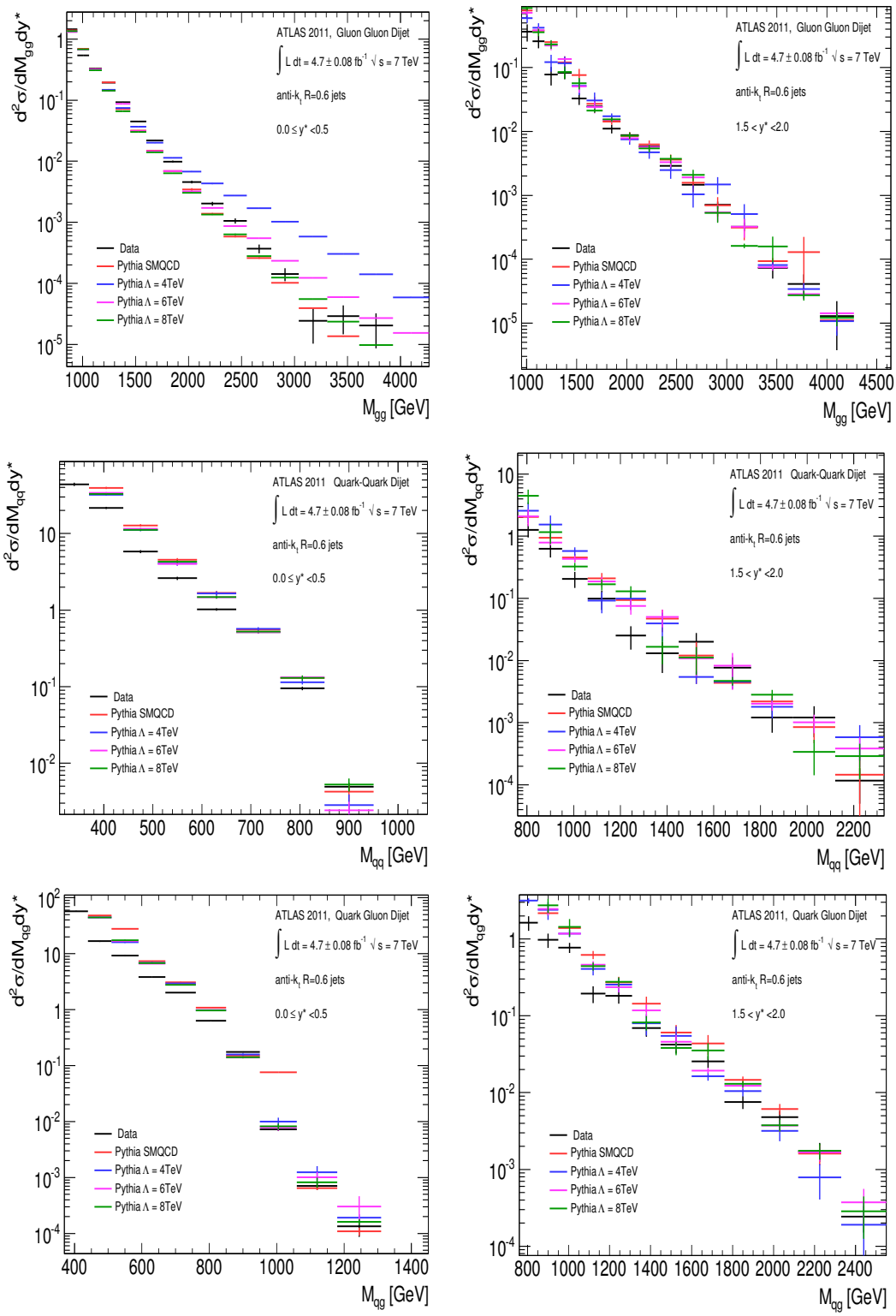


Figure 5.6: First top figures show gluon-gluon dijet invariant mass at different y^* cut at LHC in pp collision at $\sqrt{s} = 7$ TeV with the full 2011 data sample, the two figure in the middle show quark-quark dijet invariant mass at different y^* cut. Last two figure at bottom show quark-gluon dijet invariant mass at different y^* cut 2011 data (black line) is compared with SMQCD (red line) and $\Lambda = 4$ TeV (blue line), $\Lambda = 6$ TeV (pink line), $\Lambda = 8$ TeV (green line) of Contact Interaction

5.3 Purified samples of quark- and gluon-initiated jets

With the larger dataset available in 2012, it is possible to extract from data samples of jets with a very high probability of coming from quarks or from gluons. Figures 5.7, taken from [95], show the fraction of quark-initiated and gluon-initiated jets in a series of samples, among which the ones used in this analysis. It can be observed that while γ+jet or W/Z+jet have very high quark jets fractions, b+jet, bb+jet, b+2-jets and the dijet, 3- or 4-jets samples have large fractions of gluon jets. The process γ+jet has a 20% contribution from final-state gluons coming from $qq^- \to g\gamma$, and an 80% contribution from quark jets coming from $qg \to q\gamma$. The quark purity can be further enhanced by applying additional kinematical cuts. The quark-jet purity is similar for the case of the $W+$ jets process. For pure gluon samples, even though b+2-jets and three jet have good fraction on gluon jets, but their cross sections are orders of magnitude smaller than the 2-jets sample as seen in bottom Figure 5.7. Since a large number of jets is needed to build templates, the dijet sample is chosen for building the gluon jet templates.

5.4 Data Sample for the 2012 analysis

For the 2012 analysis, the full luminosity of 20.3 fb^{-1} has been used. In order to extract the purified samples, different datasets are used for the di-jet, γ+jet and W+jet selections. Events of 2→2 qq, gq, or gg scattering with Pythia8 with the AU2CT10 tune and the CTEQ10 pdf, Alpgen Pythia with P2011C tune, Pythia8 with AU2 tune and CTEQ6 pdf set are used respectively for dijet, W+jet and γ+jet Monte Carlo.

Each of the three samples has been selected with specific requirements on trigger objects, that will be described in each separate subsection. The jet selection is however common, described in the following.

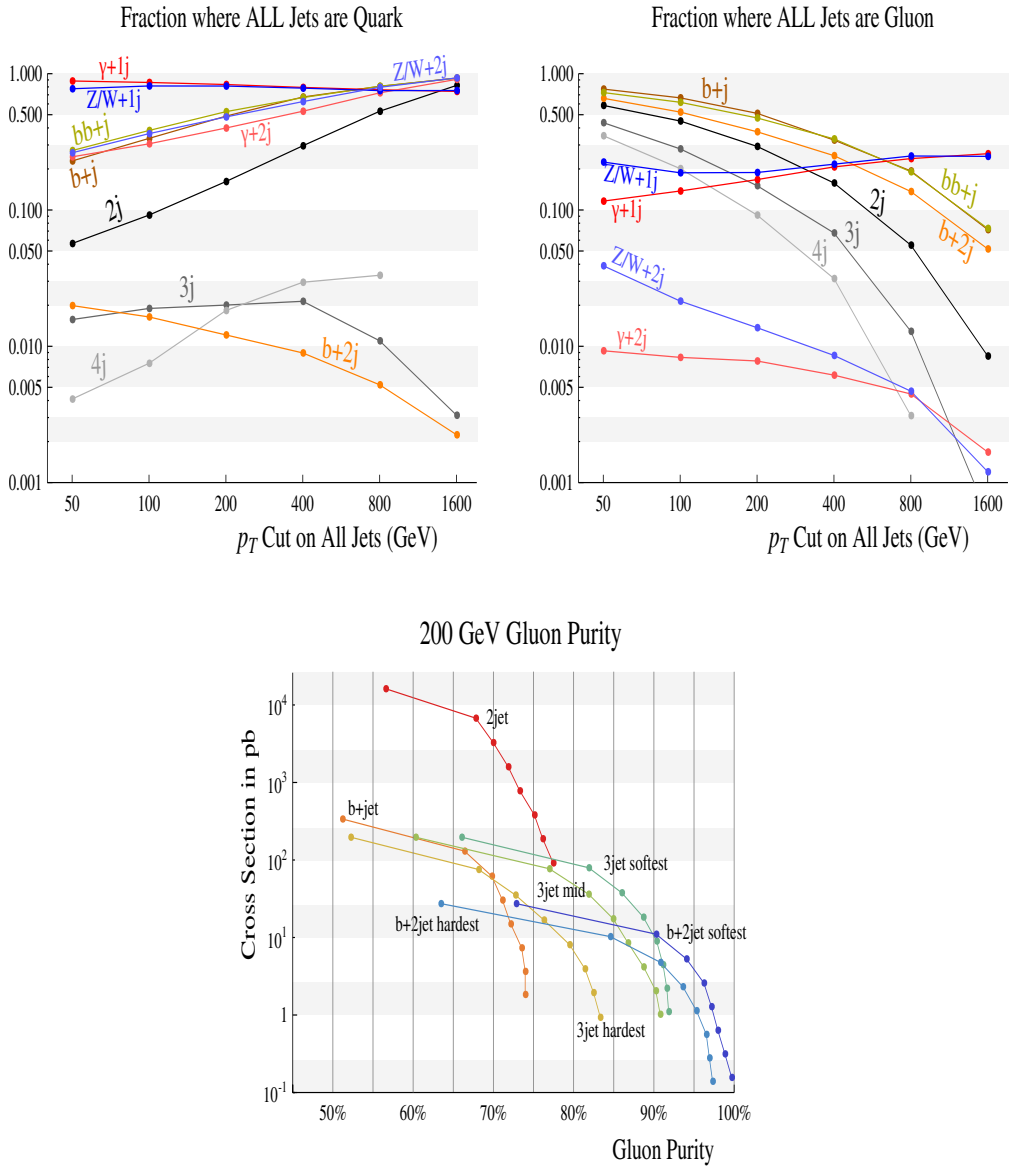


Figure 5.7: The top row shows fraction of events where all jets are quark or gluon, on a log scale. The bottom figure shows the cross section as a function of gluon purity for the different samples with a 200 GeV cut on all non-b jets [95].

5.4.1 Jet Selection

Anti-k_{T} jets with a radius parameter of 0.4 and calibrated with the LCTopo[2] method are used in this analysis. The selection requirements used for these jets are:

- $p_T > 20$ GeV
- jet does not fall in the isBadLoose nor isUgly categories
- If $p_T < 50$ GeV and $|\eta| < 2.4$. It must pass jvf[3] requirement > 0.25 to reject spurious calorimeter jets from local fluctuation in pile-up activity as well as real QCD jets originating from single pile-up interactions.
- $\Delta R_{\text{leading jet, nearest jet}} > 0.8$ with respect to the nearest jets with $p_T > 15$ GeV. The nearest jets do not need to satisfy the jvf.

The last isolation requirement is used to decreases contributions from both pile-up and QCD color-connection to nearby jets.

5.4.2 Dijets

Events classified as dijets must satisfy following selection, on top of the jet requirements:

- The leading jets must pass the trigger requirements as in Table 5.3
- The leading jet must have $p_T > 25$ GeV
- The subleading jet must have $p_T > 20$ GeV
- $\Delta\phi_{\text{leading,subleading}} > 2.5$ to veto soft radiation
- $|\eta_{2^{\text{nd}}\text{ jet}}| > |\eta_{1^{\text{st}}\text{ jet}}|$ and same z-hemisphere.

[2]LC (Local Cluster): is calibration of calorimeter cluster. It attempts to separate EM from Hadron calorimeter cluster and apply individual correction before jet reconstruction.

[3]jvf , or jet vertex fraction, is the fraction of tracks associated to the jet which come from the primary vertex

Leading Jet p_T	**Trigger**
25 GeV	EF_j15_a4tchad or EF_fj15_a4tchad [4]
36 GeV	EF_j25_a4tchad or EF_fj25_a4tchad
47 GeV	EF_j35_a4tchad or EF_fj35_a4tchad
73 GeV	EF_j55_a4tchad or EF_fj55_a4tchad_L2FS
115 GeV	EF_j80_a4tchad or EF_fj80_a4tchad
142 GeV	EF_j110_a4tchad or EF_fj110_a4tchad
174 GeV	EF_j145_a4tchad or EF_fj145_a4tchad
218 GeV	EF_j180_a4tchad or EF_fj180_a4tchad
268 GeV	EF_j220_a4tchad
392 GeV	EF_j360_a4tchad
512 GeV	EF_j460_a4tchad

Table 5.3: Leading jet p_T bins used to select the trigger used for the di-jet samples.

The reason for requiring that the two jets are in the same direction in the longitudinal plane is the increased gluon fraction of forward jets. If the leading jets is observed in the central region, the subleading jet can be anywhere. However when the leading jet is in the forward (backward) region, the subleading jet is also required to be in the same region. Requiring a momentum-imbalance along the beam axis selects events more likely to have originated from a quark and gluon pair. Moreover, the presence of a gluon in the initial state gives more chances to have a gluon in the final state.

5.4.3 γ +jet

The jets requirements for the γ + jet sample are:

- The event must pass one of the photon triggers listed in table 5.4. The p_T binning was chosen if the trigger is fully efficient above p_{T}.

- Ignore all jets overlapping with the leading photon with $\Delta R < 0.2$. This

[4] **EF fjX a4tchad:** requires at least a jet with transverse energy (E_T) above X GeV at the EF at forward region. a4tchad means that jets are reconstructed using the anti-kt algorithm with topological cell energy clusters as signal input with a radius of R=0.4

cut provides to make better photon reconstruction.

- | Leading photon η| < 1.37 or 1.52 < | Leading photon η| < 2.5. Out of this region, γ efficiency and rejection are worse because of large amount of passive material of EM calorimeter.

- Isolation: hadronic energy in a cone of ΔR < 0.4 around the photon must be < 3 GeV to remove large background (e.g $\pi^0 \to \gamma\gamma$) and help to reduce fragmentation photons. Isolation also provides to decrease soft-jet coming from pile-up and underlying event.

- The leading photon must satisfy quality criteria to remove bad photons affected by dead material for LAr cluster and masked cells. The leading photon also must pass from cleaning criteria to suppress cluster with large energy contribution from bad cells.

- $|\Delta\phi_{Leading\ jet,Leading\ photon}| > 2.9$ to suppress initial state radiation

- Subleading jet p_T < max(40 GeV, 0.3 $\times p_{T\text{leading }\gamma}$)

Leading Photon p_{T}	**Trigger**
25 GeV	EF_g20_loose
45 GeV	EF_g40_loose
65 GeV	EF_g60_loose
85 GeV	EF_g80_loose
105 GeV	EF_g100_loose
125 GeV	EF_g120_loose

Table 5.4: Leading photon p_T bins used to select the trigger used for the gamma-based samples.

5.4.4 W+jet

Events falling in this category must satisfy the following requirements:

- Event must pass at least one of the two muon triggers EF_mu24i_tight and EF_mu36i_tight[5])

[5]EF_muXi_tight: selects events from a single muon trigger requiring at least one isolated

- Muons are required to pass the following criteria
  - 'Tight'[6] muon identification or StandAlone[7] are combined [96].
  - $p_{\mathrm{T}} > 10$ GeV
  - $|\eta| < 2.7$
  - Track isolation : $(\sum p_{\mathrm{T}}$ in cone of 0.2$)/p_{\mathrm{T}} < 0.1$
  - Calo. isolation : $(\sum E_T$ in cone of 0.3$)/E_T < 0.14$

In order to reduce large background from multijet production, muons are isolated from neighbouring track using track isolation cut and also it is isolated from other calorimeter energy depositions to correct piel-up contributions. Figure 5.8 shows the quark-jets, gluon-jets and heavy quarks-jets contributions from the W+jet sample in MC. The fraction of quark jets is almost constant to ∼87% as a function of transverse momentum over the whole η range. The bottom-jet contribution is lower than 0.1%. While the charm-jet fraction is more or less similar to the gluon-jet fraction at lower p_T, it does decrease for larger transverse momenta.

The fraction of the quark, gluon and heavy quark jets for the dijet and γ+jet(bottom) at $|\eta|$ <0.8 and 1.2 < $|\eta|$ <2.1 are shown in Figures 5.9. The dijet sample has 70% gluon-like jets at low p_T and decreases with increasing p_T. It becomes mostly quark-like at higher p_T at higher η region. The $\gamma + jet$ sample has much more quark-jets, >%70. There are significant differences in the predicted fractions by Pythia 8 and Herwig++.

5.5 Variables used for quark-gluon separation

Quark and gluon jets have different properties because of their different parton-showering processes, leading to narrower jets in the case they are quark-

muon with p_{T}> X GeV GeV where isolation criterion, "tight" is made with inner detector tracks.

[6]Tight cut: μ from W,Z. A good quality track from a combined fit of the hits in the tracker and muon system

[7]StandAlone: Information comes from only muon spectrum.

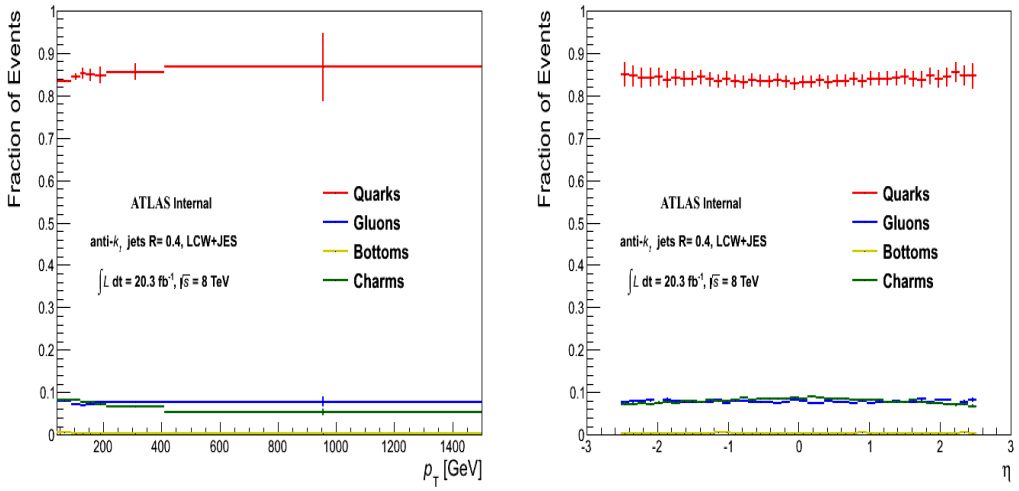


Figure 5.8: The fraction of events where the leading jet and η are a quark-jet, a gluon-jet or heavy partons in the W+jet sample, taken from MC generators.

initiated [95]. According to the proton PDFs, initial-state gluons have a lower x than initial-state quarks before the hard scattering, and that also results in a higher abundance of final-state gluon jets at lower p_T. The variables described below have shown a good separation power between these kinds of jets, in the following studies: [97] [98] [99] [100].

Number of calorimeter clusters in the jet:

$$n_{cal} = \sum_{\text{const}\in\text{jet}} \tag{5.3}$$

E_{T} weighted width of the jets:

$$w = \frac{\sum_{\text{const}\in\text{jet}} E_{\text{T,const}} \Delta R_{\text{const,jet}}}{\sum_{\text{const}\in\text{jet}} E_{\text{T,const}}} \tag{5.4}$$

Fraction of energy carried by the largest energy constituent:

$$f^{largest} = \frac{E_{\text{largestconst}}}{E_{\text{jet}}} \tag{5.5}$$

Two point energy correlation function:

$$C_\beta = \frac{\sum_{i,j\in\text{jet}} E_{\text{T},i} E_{\text{T},j} (\Delta R_{i,j})^\beta}{\sum_{i\in\text{jet}} E^2_{\text{T},i}} \tag{5.6}$$

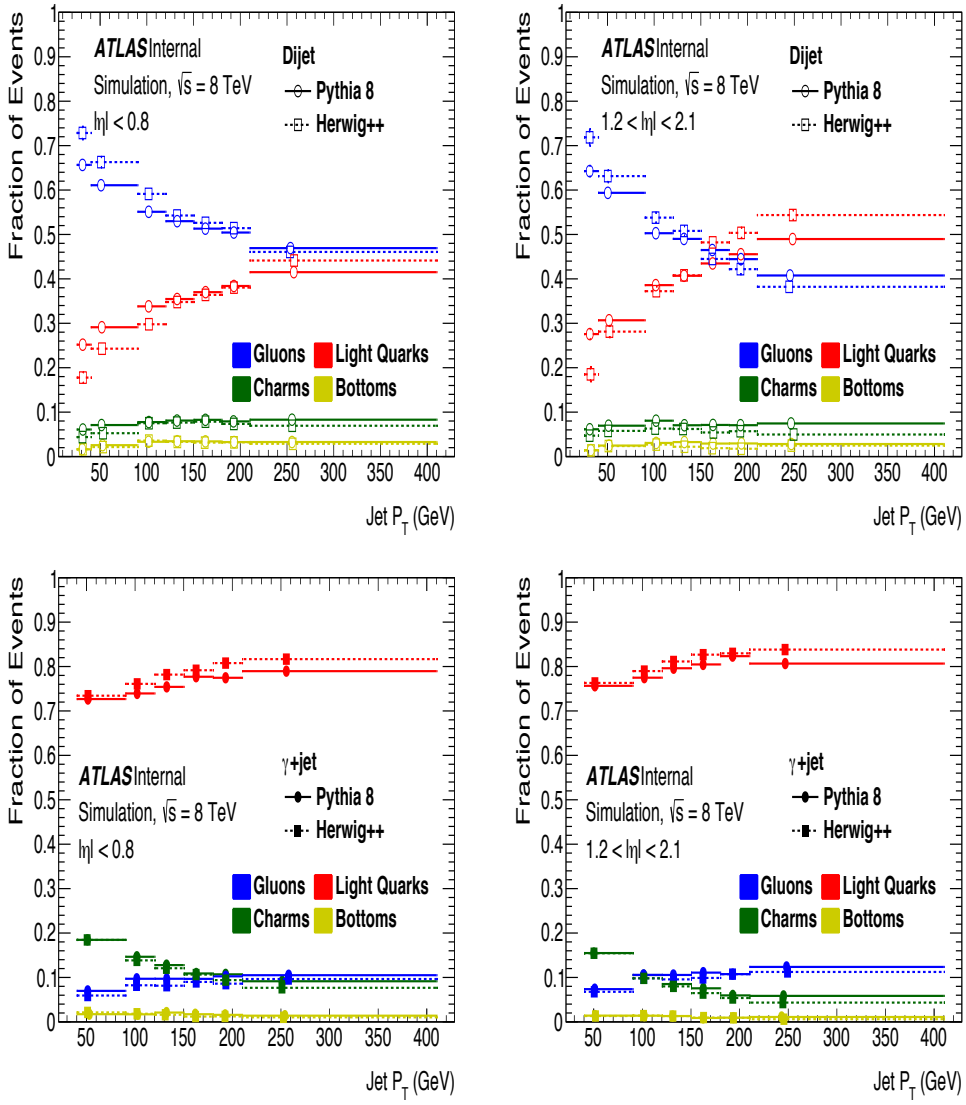


Figure 5.9: The fraction of events where the leading jet is a quark or a gluon in the dijet sample(top), and γ+jet(bottom) taken from Herwig++ and Pythia 8 samples at $|\eta|$ <0.8 and 1.2 < $|\eta|$ <2.1.

The various qualities of β (β=0.1,0.2,0.5,1,2) are applied to find better discrimination between quark and gluon jets. It is observed that β=0.2 gives much better diversity than other quantities does. These variables can be computed from tracks or from calorimeter quantities. Tracks are required to satisfy the following criteria:

- $p_{\mathrm{T}} > 1$ GeV
- Number of pixel hits ≥ 1
- Number of SCT hits ≥ 6
- $p_{\mathrm{T}} > 1$ GeV
- $|d_0|$ with respect to primary vertex $<$ 1mm
- $|z_0 \sin(\theta)|$ with respect to primary vertex $<$ 1mm

and are assigned to the jets using a ghost-matching technique[8] [101]

On the other hand, there is no additional requirement on the calorimeter clusters for calorimeter-based variables, apart from the cluster being used to be part of the jet. Figures 5.10 and 5.11 show the track width distribution versus the number of tracks and calorimeter width distribution versus the number of constituents for quark and gluon jets, from W+jet Monte Carlo events. It can be seen that, as gluons on average produce more splitting than quarks in the parton evolution, gluon jets are broader, and have more tracks.

5.6 Template extraction from two event samples

The jet-shape templates are extracted from two samples, dijet and W+jet events. Templates are built by reweighting the events according to the expected fractions of quarks and gluons. The shape of a template distribution for a

[8]matching of tracks to jets in ΔR

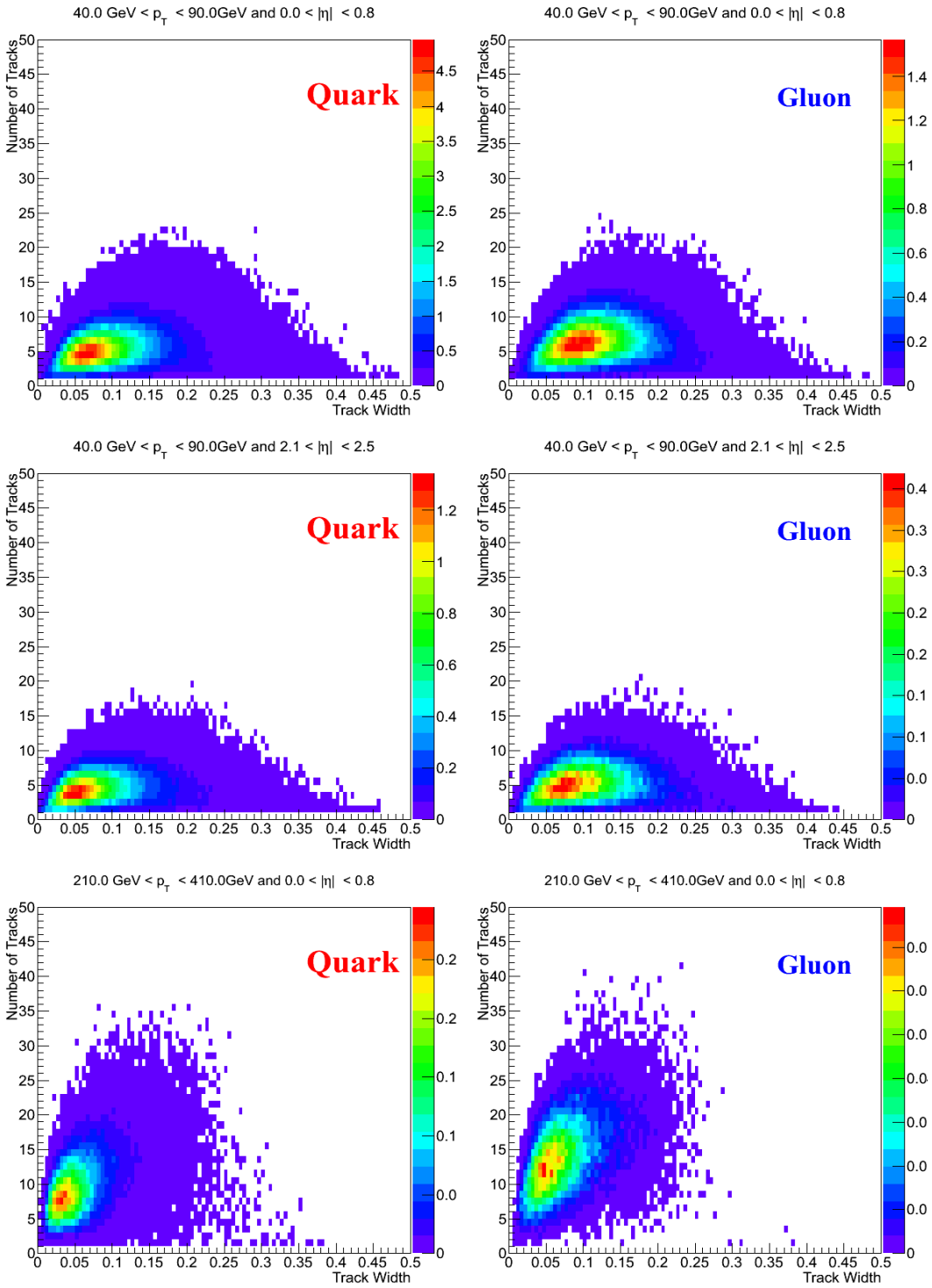


Figure 5.10: 2D plots showing the number of tracks versus the track width in different $|\eta|$ and p_T range. The left plots refer to quark jets and the right ones are to gluon jets.

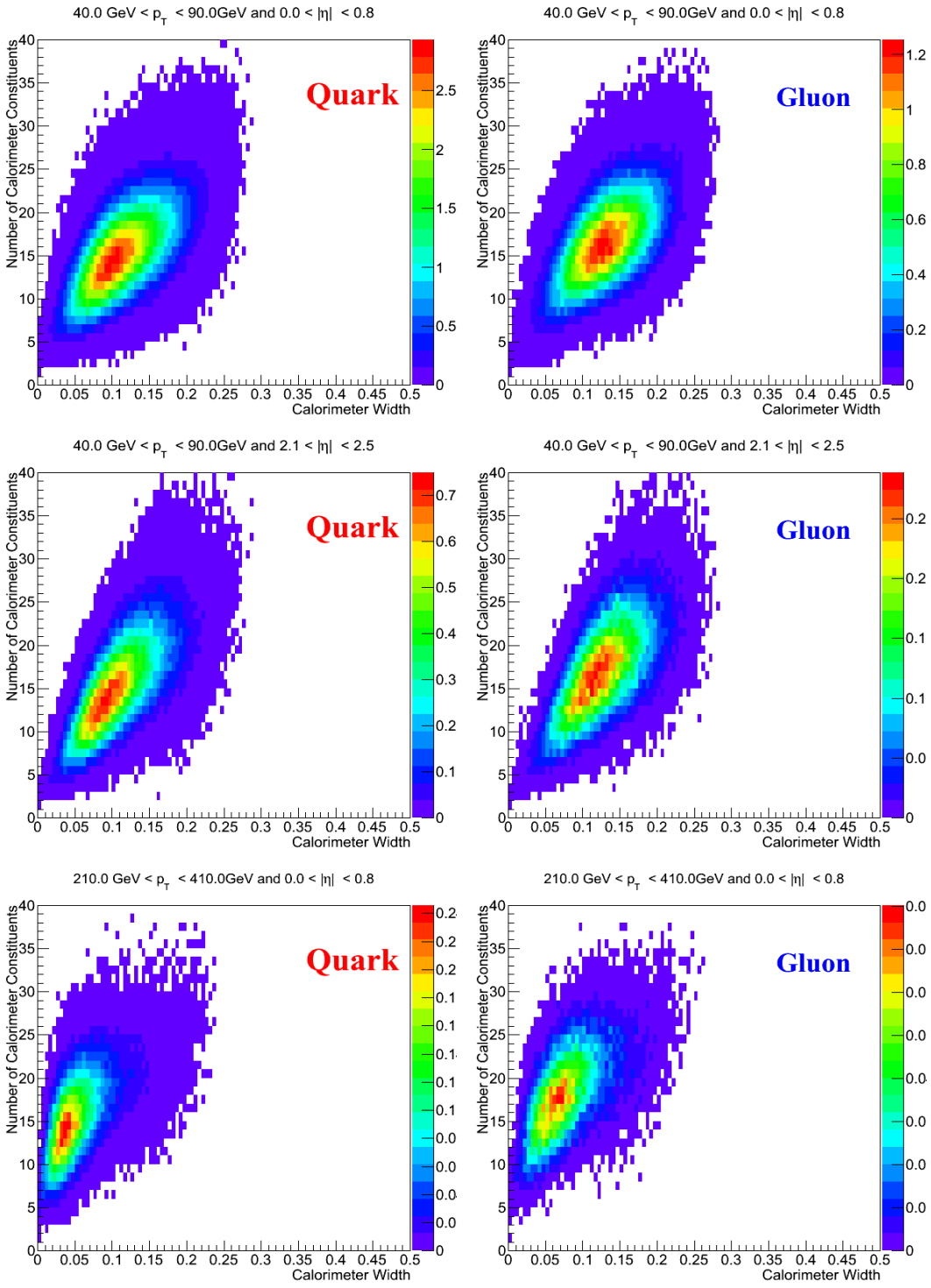


Figure 5.11: 2D plots showing the number of calorimeter constituents versus the calorimeter width in different $|\eta|$ and p_T range. The left plots refer to quark jets and the right ones are to gluon jets.

variable is a linear combination of the shape for quark jets, weighted by the fraction of quarks, and the gluon-jet shape, weighted by the amount of gluons in that sample. In addition, charm and bottom contamination is weighting by the expected fraction of these quarks, and subtracted using MC sample.

For two independent samples (the dijet and W+jet), the relations between the inclusive shapes are described by the following equation:

$$\begin{pmatrix} p_{\text{dijet},i} \\ p_{W+jet,i} \end{pmatrix} = \begin{pmatrix} f_{\text{dijet},q} & f_{\text{dijet},g} \\ f_{W+jet,q} & f_{W+jet,g} \end{pmatrix} \begin{pmatrix} p_{q,i} \\ p_{g,i} \end{pmatrix} + \begin{pmatrix} f_{\text{dijet},hf} p_{\text{dijet},hf,i} \\ f_{W+jet,hf} p_{W+jet,hf,i} \end{pmatrix} \tag{5.7}$$

where the subscript q and g means quarks or gluons. The subscript i represents to the ith bin of the distribution histogram. $f_{sample,q(g)}$ is the fractional contribution of gluons and quarks in MC to each sample. $p_{q,i}$ or $p_{g,i}$ are the value of the normalised pure distributions in bin i. On the left hand side of formula, the value of the normalised inclusive distributions in bin i are $p_{\text{dijet},i}$ and $p_{W+jet,i}$. The contributions of charm and bottom quarks in samples(dijet or W+jet) are given in a similar notation, and indicated by hf. These shapes are calculated independently for each p_T and $|\eta|$ bins.

There are two unknowns, $p_{q,i}$ and $p_{g,i}$, in Equation 5.7. Since it is a linear system of two equations, this system can be solved analytically. When the quark/gluon fractions are similar in both samples, the matrix in equation 5.7 has a very small determinant. In this case, the matrices with determinants close to zero tend to be unstable to inversion, which causes large fluctuations in the solution. To avoid this issue, kinematic cuts have been applied to make the dijet sample have as much gluon-like jets as possible as described in subsection 5.4.2.

Figures 5.12 and 5.13 illustrate templates extracted from data belonging

to the dijet and W+jet samples. The variables used are jet track width, number of tracks, calorimeter width and number of calorimeter clusters, for different $|\eta|$ bins. The discrimination between quark and gluon jets is much stronger for track width and calorimeter width comparing to other variables. The extracted templates(data)-Monte Carlo agreement for these variables is fairly good for number of calorimeter constituents, track width and calorimeter width; however there is not a good agreement for gluon-like jets in the number of calorimeter constituents.

5.7 Template extraction from three event samples

In the previous case, quark-jet "pure" templates have been extracted from the W + jets sample; however the statistics is not very large, resulting in some fluctuations for the data-driven templates. The γ+jet sample has been added in order to increase statistics.

Since the quark and gluon fractions are different between the W + jets and the γ + jets samples, the extraction of the templates from three samples has to be extended using a more complicated equation. For three independent samples (in this case dijet, γ and W+jet), equation 5.7 gets modified in the following:

$$\begin{pmatrix} p'_{dijet,i} \\ p'_{\gamma+jet,i} \\ p'_{W+jet,i} \end{pmatrix} = \begin{pmatrix} f_{dijet,q} & f_{dijet,g} \\ f_{\gamma+jet,q} & f_{\gamma+jet,g} \\ f_{W+jet,q} & f_{W+jet,g} \end{pmatrix} \begin{pmatrix} p_{q,i} \\ p_{g,i} \end{pmatrix} \tag{5.8}$$

where i is the bin number in the distribution, $p_{j,i}$ is the number of events in bin i from data sample j, after correction for heavy flavors and fakes.

Figures 5.14 through 5.16 show templates for several variables and various p_T and η bins, extracted from the dijet, W+jet and gamma+jet samples. The agreement between data and Monte Carlo improves after addition of the third sample.

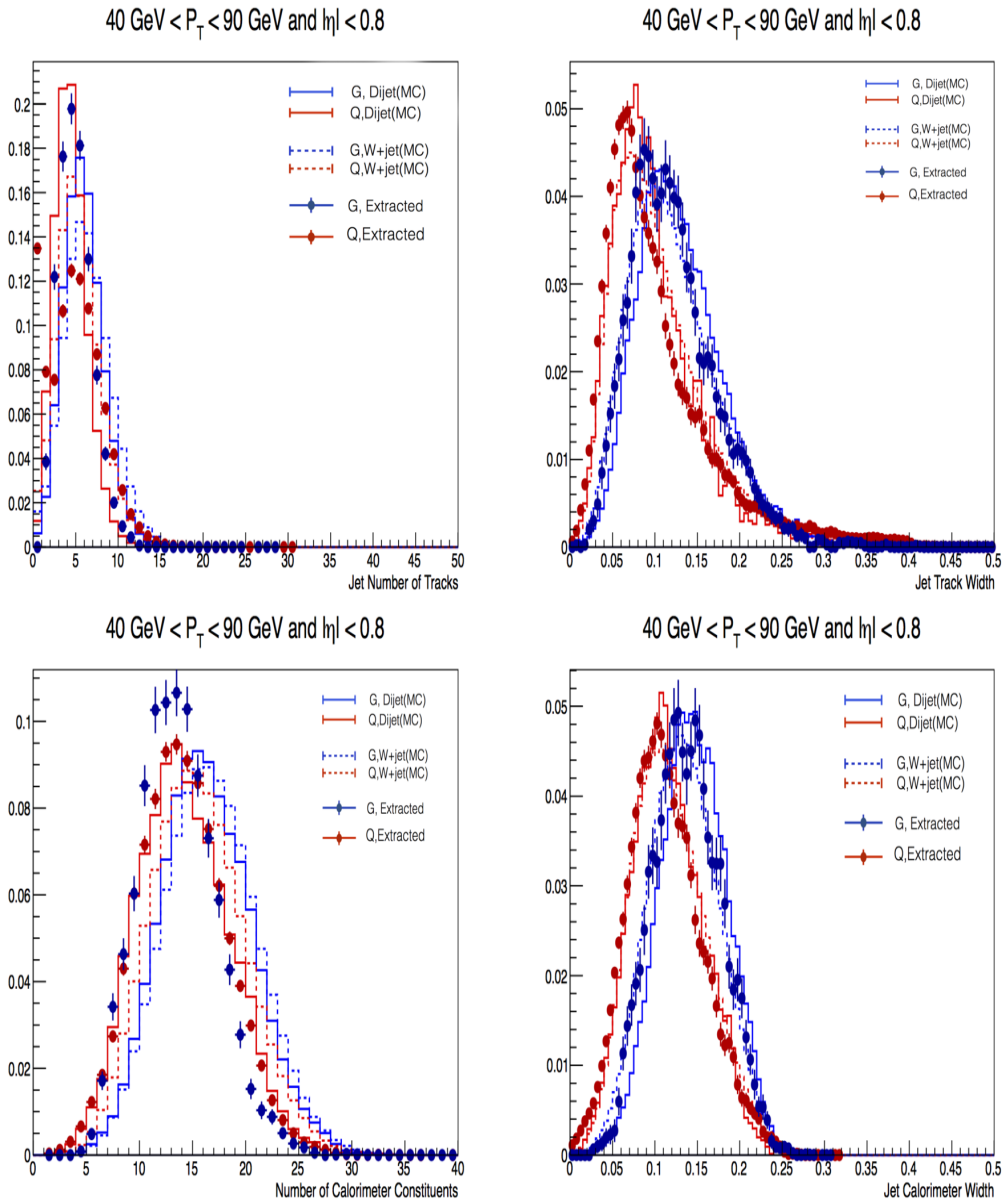


Figure 5.12: Templates extracted from dijet and W+jet samples for jet track width, number of track, calorimeter width, number of calorimeter at $40 < p_T < 90$ GeV and $|\eta| < 0.8$. G is gluon and Q is quark.

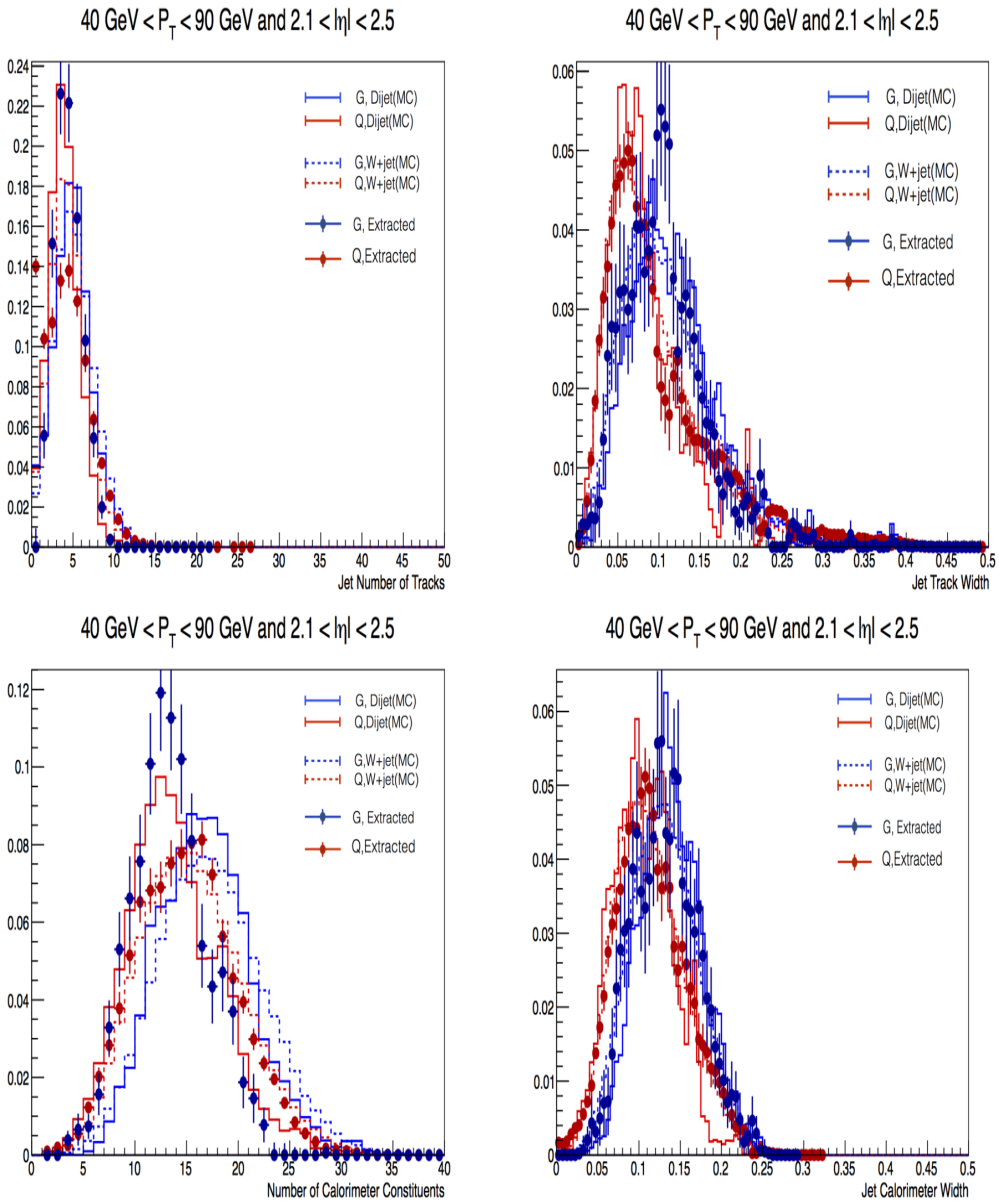


Figure 5.13: Templates extracted from dijet and W+jet samples for jet track width, number of track, calorimeter width, number of calorimeter at $40 < p_T < 90$ GeV and $2.1 < |\eta| < 2.5$.G is gluon and Q is quark.

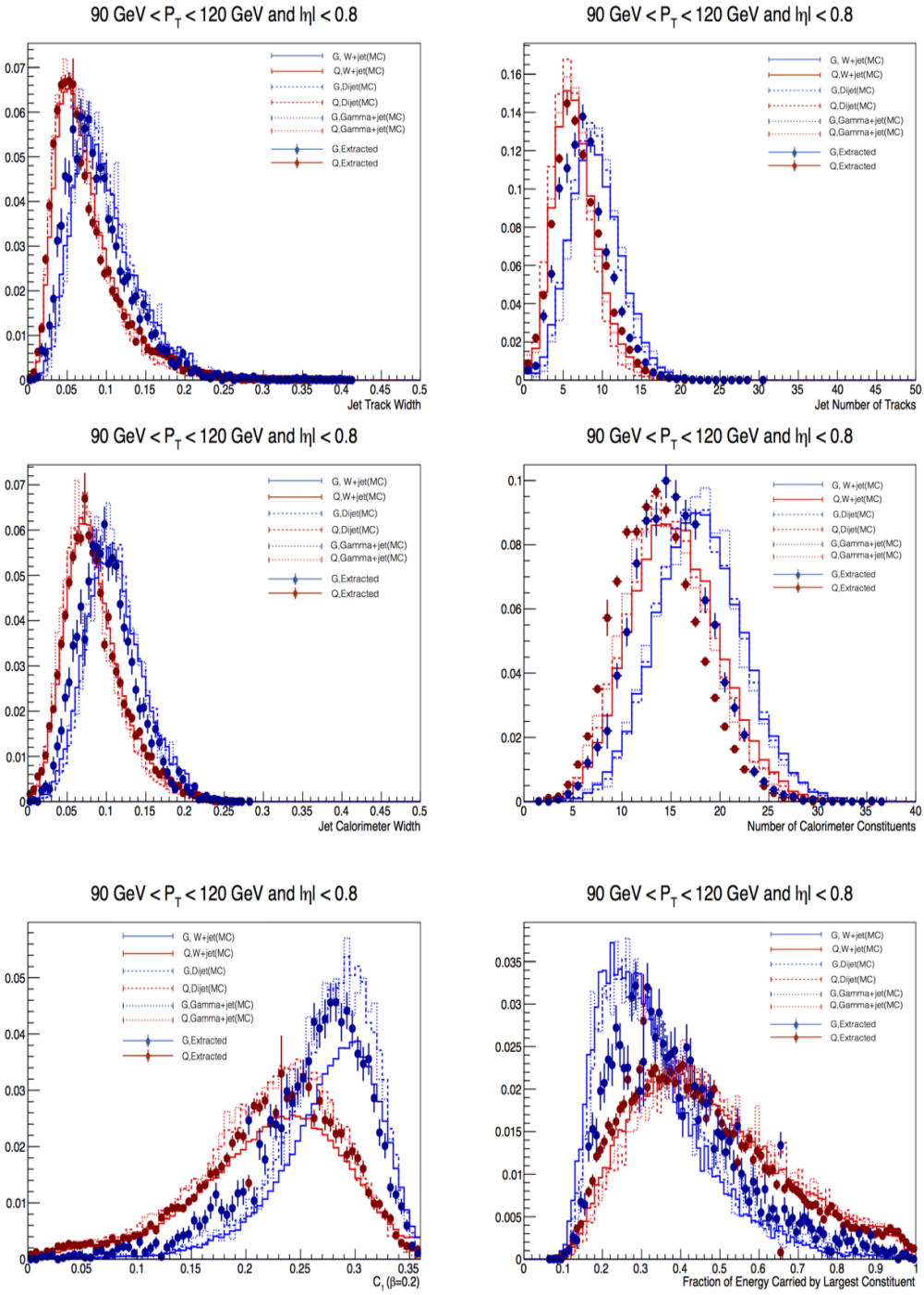


Figure 5.14: Templates extracted from the dijet, W+jet and γ+jet samples for jet track width, number of track, calorimeter width, number of calorimeter, C_β called as C_1 (β=0.2), fraction of Energy carried by largest constituents at $90 < p_T < 1200$ GeV and $|\eta| < 0.8$.G is gluon and Q is quark.

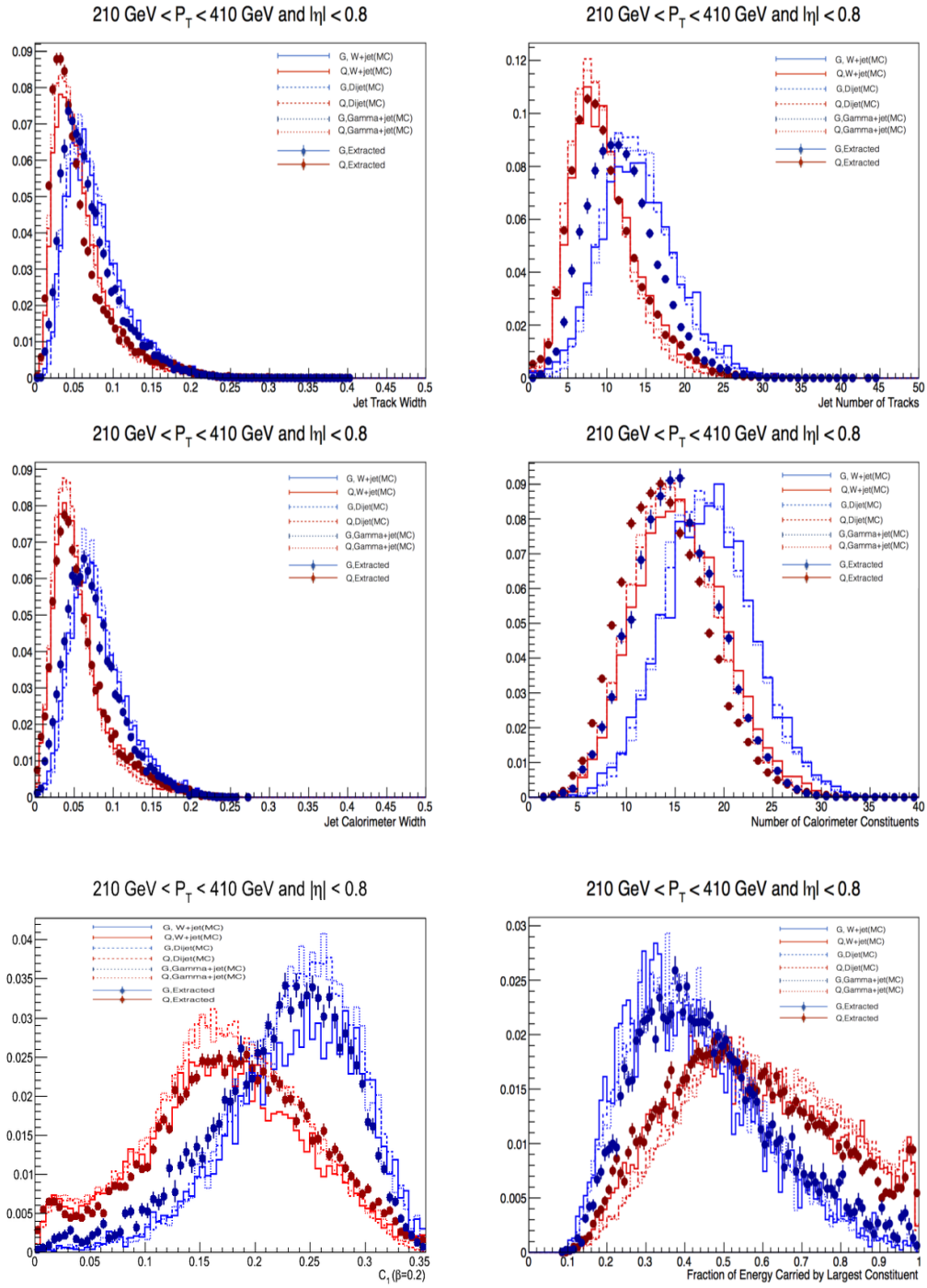


Figure 5.15: Templates extracted from dijet, W+jet and γ+jet samples for jet track width, number of track, calorimeter width, number of calorimeter, C_β called as C_1 (β=0.2), fraction of Energy carried by largest constituents at $210 < p_T < 410$ GeV and $|\eta| < 0.8$.G is gluon and Q is quark.

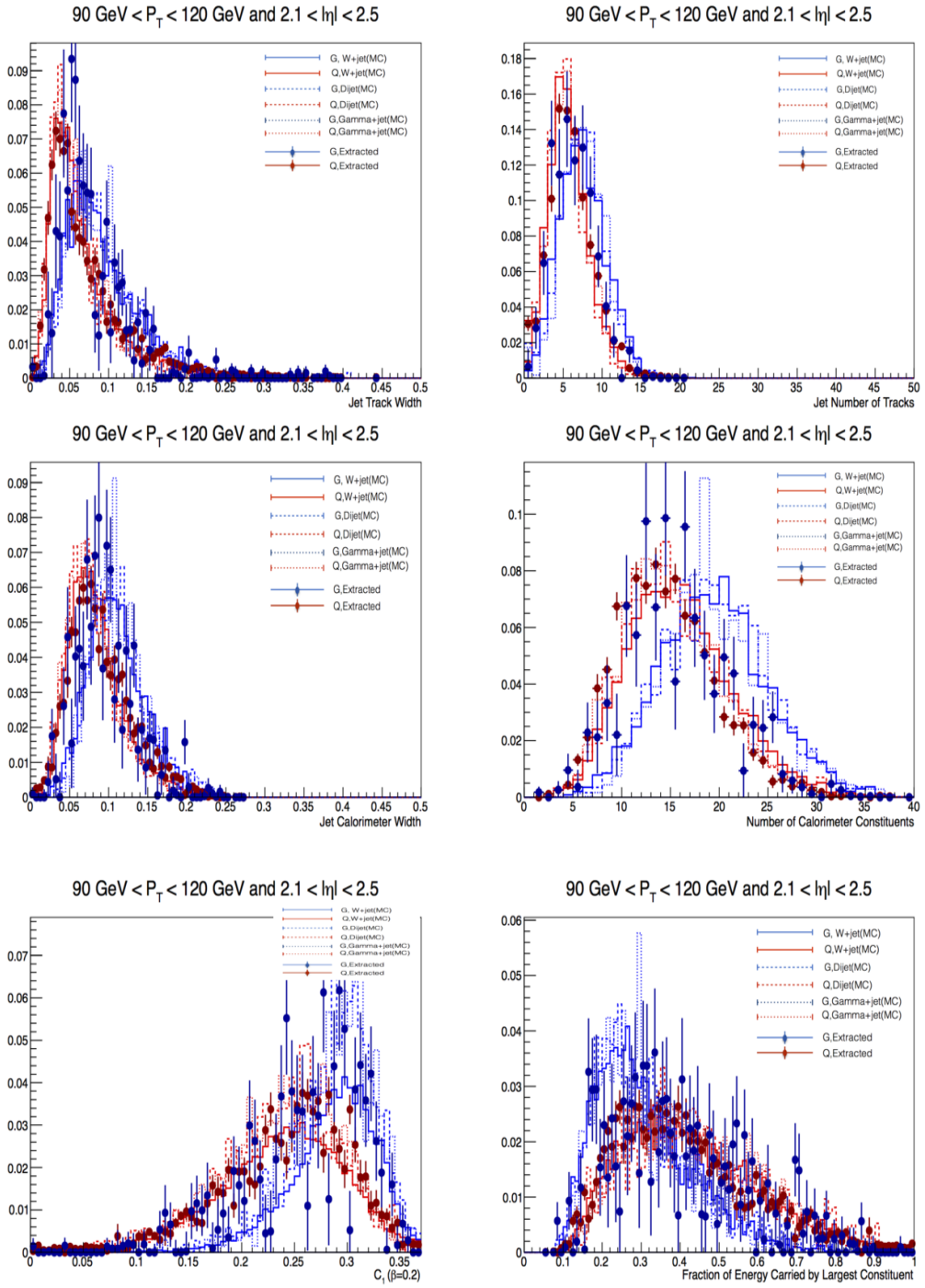


Figure 5.16: Templates extracted from dijet, W+jet and γ+jet samples for jet track width, number of track, calorimeter width, number of calorimeter,C_1, fraction of Energy carried by largest constituents at $90 < p_T < 120$ GeV and $2.1 < |\eta| < 2.5$.G is gluon and Q is quark.

5.8 Quark-gluon separation performance

The variables described in the previous sections have been used to perform a discrimination between quark and gluon jets in the pure samples obtained in from data. By varying the cut on each variable, curves of quark acceptance vs gluon rejection power can be produced in various kinematical regions. Figure 5.17 shows the behaviour of the tagger in two intervals of transverse momentum, and also for the forward rapidity. It can be seen that the energy-energy correlation variable C_β shows the best separation between quark and gluon jets, however statistical uncertainties are quite large, resulting in possible fluctuations of the curve.

5.9 Conclusions for the quark-gluon separation study

In the analysis about contact interaction in 2011, we have investigated the possibility to discriminate between jet production as described by QCD and the same process in the presence of Contact Interaction. While the shapes of the various distributions are slightly different for quark- and gluon-tagged jets, and some additional separation could be gained by separating the samples, the main differences occur at very large values of jet transverse momenta and dijet invariant masses, where statistical fluctuations are the highest. The quark-gluon tagger was not built well and it did not help to observe obvious result on the analysis of contact interactions in the 2011 dataset as desired. In order to improve the tagger, the increased luminosity of the 2012 dataset was used, and additional variables were considered. To built the tagger, templates have been extracted from data, using two or three samples of relatively pure jets. For a fixed quark-like jet acceptance of 70%, the variables C_β (fraction of energy carried by the largest constituent), number of track, calorimeter width and track width are able to reject more than 40% of the gluon-like jets,

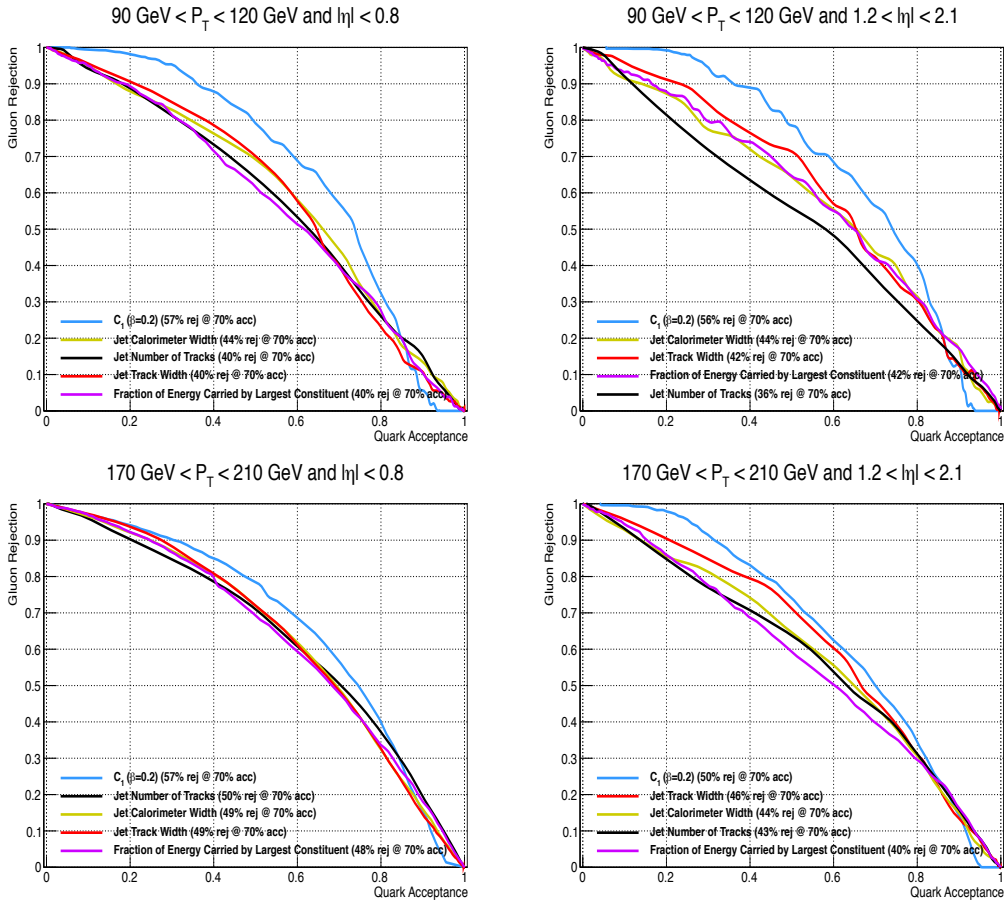


Figure 5.17: Gluon rejection versus quark Acceptance in data for likelihoods built from extracted templates from "three samples" at 90 GeV $< p_T <$ 120 GeV (top) and at 170 GeV $< p_T <$ 210 GeV (bottom), at $|\eta| < 0.8$ (left) and , $1.2 < |\eta| < 2.1$ (right). Dijets, W+jet and γ+jet samples were used for the extraction.

while number of calorimeter constituents only ∼20% of gluon-like jets. The purified samples used were dijet, γ+jet and W+jet. At higher $|\eta|$, statistical uncertainties become large, so the comparison with data becomes more difficult. The number of calorimeter constituents shows the largest discrepancy between extracted templates and Monte Carlo. If I had time, I would complete 2012 quark-gluon separation tagger with "Discrimination of Light Quark and Gluon Jets" study group and continue to work on analysis of contact interactions. I would improve exclusion limits on the compositeness scale for quark CI ranges.

Chapter 6

The Run Time Tester (RTT)

6.1 The Run Time Tester (RTT)

Every night the ATLAS software is recompiled using several different "builds" (about 25). A build is a term referring to the compilation, linking and making of libraries for a determined set of package tags. Examples of parameters than can vary between builds include the package tags used, the operating system targeted and the compiler version as well as the compiler options selected.

The Run Time Tester, also known as RTT [102], is a Python-coded framework [103] used to test the ATLAS software. It is a facility for running a set of Athena jobs [104], performing actions on the job results, and publishing information about these results. The RTT monitoring should ensure the successful completion of each individual job submitted by a developer. It has the option of being started via an automatic "cron" job which is the scheduled task itself, or locally in the shell. Jobs run on 70 computers for durations ranging from minutes to hours depending on the job size. A xml file [1]is used to select the configuration information for each RTT release. The xml file can specify branch/platform, packages, and also which files to keep (for instance root or log files).

An RTT test run consists of some steps like job submission, job check-

[1]xml file is a markup language that defines a set of rules for encoding documents in a format.

ing, other post job activities, and checking the result (histograms, summary). CMT[2] [105] is used in RTT to establish the connection between the release information and the actual libraries and executable. Test jobs are defined within an xml unified test configuration file and then jobs are submitted to the computing farm. The structure of RTT processing is shown in Figure 6.1.

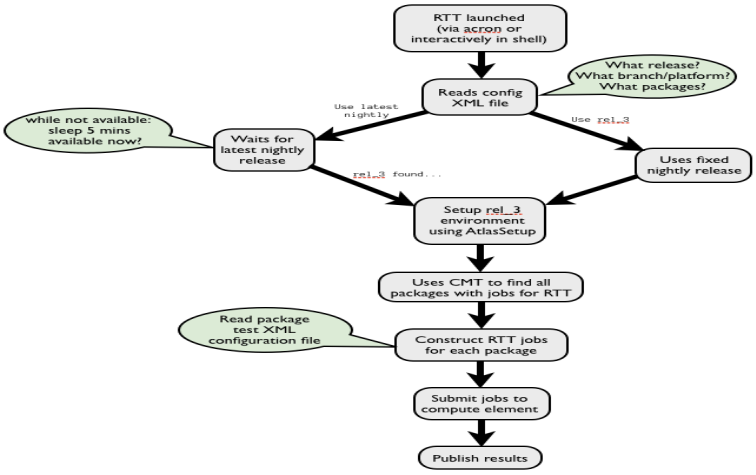


Figure 6.1: The RTT execution from starting steps to end [102].

Multiple jobs can be run using a chain in a single xml to indicate that more than one independent files can simultaneously run in parallel, and then their input files can be transferred to subsequent jobs, as illustrated in Figure 6.2.

6.2 Personal contribution to RTT running

As a service work to obtain qualification as an ATLAS author, I was involved in a shift task to run and monitor the RTT on a daily basis for events reconstructed by the JetTauETMiss stream, e.g. selected by a trigger requiring a jet, a tau lepton or missing transverse energy. In a daily RTT shift, the production process of a dataset (the RDO, ESD, AOD and D3PD formats) must be tested comparing the output of the new dataset with that of a ref-

[2]The configuration and build program used to build ATLAS code.

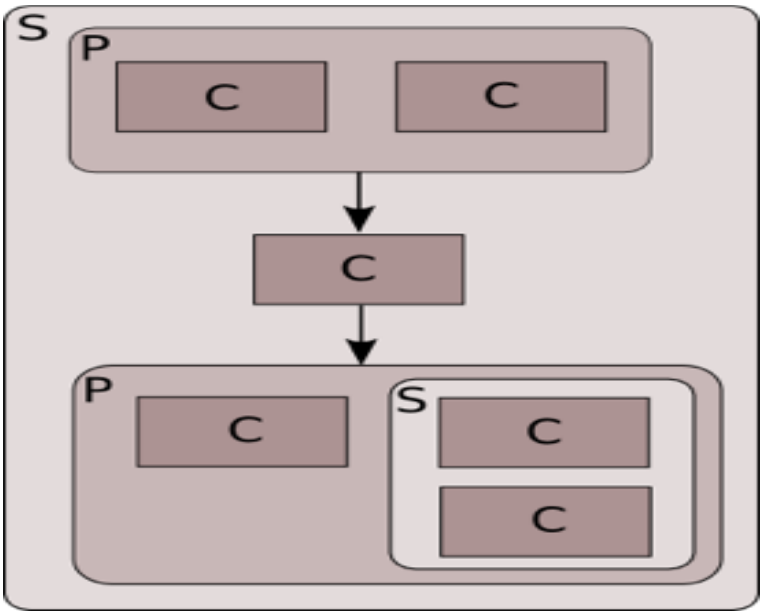


Figure 6.2: Schematic of a job chain showing S(sequential), P(parallel) and C(chain minder) jobs

erence file. The aim of this test is to ensure that no information is lost or corrupted in the new dataset. All information is compared branch by branch, variable-by-variable, event-by-event, and jet-by-jet.

To produce a dataset in the D3PD format, the following production steps are needed: RAW $\rightarrow$ ESD $\rightarrow$ AOD $\rightarrow$ D3PD for data, and RDO $\rightarrow$ ESD $\rightarrow$ AOD $\rightarrow$ D3PD for Monte Carlo [106]. Through the xml file, the shifter checks in each step that all information is transferred into the new dataset. The test is made first locally with a small number of events, and then same process is repeated running on larger files using an automatic job. I have been involved in the production and test of data after the introduction of new jet reconstruction variables, such as AntiKtXLCTopoJets or AntiKtXTopoEMjets.

Chapter 7

Conclusions

This thesis describes two analyses: the measurement of the inclusive jet cross-section using low μ runs, and the study of quark-gluon jet discrimination. Both analyses use the data taken at a centre-of- mass energy of 8 TeV in 2012.

In the first part of this thesis, the inclusive jet differential cross-section is measured using low instantaneous luminosity conditions (a no pile-up and an intermediate-μ run) as a function of the jet transverse momentum, in bins of $|y|$, from 20 GeV to 134 TeV. The combination of this measurement with the high-momentum one performed on high-luminosity runs spans two orders of magnitude in transverse momentum by my colleague who is Gagik Vardanyan, beyond the energetic reach of any previous experiment. A significant decrease of the cross-section uncertainties has been observed compared to the 2011 jet cross-section measurement. The measurement has been performed for two values of the jet radius parameters: 0.4 and 0.6, and the systematic uncertainty for R = 0.4 is in general larger than the corresponding one for R=0.6, which has better agreement with data.

In the second part of the thesis, we have investigated the possibility to discriminate between jets originated by quarks and by gluons, with the aim of applying this tool to the search for contact interactions. This hypothetical model modifies the distribution of transverse momentum and dijet mass, in particular for large momentum transfers. Since the same effect could be

due to harder PDF sets, discriminating between quark and gluon jets could help disentangling new physics from a QCD effects for this analysis. Since the discrimination capability observed in 2011 data is too small to justify the additional complication of using this tool, the quark-gluon discrimination tool has been improved for the 2012 dataset, where the four variables track width, track multiplicity, calorimeter width, number of constituents and Fraction of energy carried by the largest energy constituent are used. It has been observed that the number of calorimeter constituents shows the largest discrepancy between data and Monte Carlo, but the large statistics collected in 2012 allowed the use of data-driven templates to extract these quantities. The result is an improved tool for quark-gluon discrimination under 2012 conditions.

Bibliography

[1] C. Burgess and P. Moore, *The standard model: A primer*. Cambridge University Press, 2006.

[2] I. Aitchison and A. Hey, *Gauge Theories in Particle Physics*, New York and London: Taylor and Francis Group, **vol 1** (2003) p. 53–73.

[3] I. Antoniadis, J. Engelen, and G. Isidori, *The W and Z particles: a personal recollection*, CERN Courier **44** .

[4] *The Higgs boson*, Website. `http://home.web.cern.ch/topics/higgs-boson`. 2015.

[5] D. Neuenschwander, *Emmy Noether's Wonderful Theorem*. The Johns Hopkins University Press, Baltimore, Published 2011.

[6] *The Standart Model*, Website. `http://www.physik.uzh.ch/groups/serra/StandardModel.html`. April 25, 2015.

[7] G. Rajasekaran, *Fermi and the Theory of Weak Interactions*, Resonance J. Sci. Educ. **19** (2014) no. 1, 18–44, `arXiv:1403.3309 [physics.hist-ph]`.

[8] A. Grozin, *Lectures on QED and QCD*, in *3rd Dubna International Advanced School of Theoretical Physics Dubna, Russia, January 29-February 6, 2005*. 2005. `arXiv:hep-ph/0508242 [hep-ph]`.

[9] B. Kopeliovich and A. Rezaeian, *Applied High Energy QCD*, Int.J.Mod.Phys. **E18** (2009) 1629–1696, arXiv:0811.2024 [hep-ph].

[10] D. Roberts Craig, *Strong QCD and Dyson-Schwinger Equations*, arXiv:1203.5341 [nucl-th].

[11] M. Bilenky and J. Hosek, *Glashow-Weinberg-Salam Theory of Electroweak interaction and the neutral current*, Phys. Rept. **vol 90** (1982) p. 73–157. 10.1016/0370-1573(82)90016-3.

[12] R. T. Chislett, *Studies of hadronic decays of high transverse momentum W and Z bosons with the ATLAS detector at the LHC*. PhD thesis, University College London, 2014. http://www.hep.ucl.ac.uk/theses/Chislett_thesis.pdf.

[13] S. Bethke, *Experimental tests of asymptotic freedom*, Nucl. Phys. Proc. Suppl. **54A** (1997) 314–326, arXiv:hep-ex/9609014 [hep-ex].

[14] R. E. Marshak, *Conceptual foundations of modern particle physics*. 1993.

[15] B. Martin, *Nuclear and Particle Physics: An Introduction*, p. 200. Wiley, 2009.

[16] R. Placakyte, *Parton Distribution Functions*, in *Proceedings, 31st International Conference on Physics in collisions (PIC 2011)*. 2011. arXiv:1111.5452 [hep-ph].

[17] D. A. Martin, *Proton structure, Partons, QCD, DGLAP and beyond*, Acta Phys. Polon. **B39** (2008) 2025–2062, arXiv:0802.0161 [hep-ph].

[18] D. Diakonov, *QCD scattering: from DGLAP to BFKL*, CERN Courier, (July, 2010) p. 24.

[19] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. Yuan, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[20] ZEUS, H1 Collaboration, F. D. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA*, JHEP **01** (2010) 109, arXiv:0911.0884 [hep-ex].

[21] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph].

[22] G. Watt, *MSTW PDFs and impact of PDFs on cross sections at Tevatron and LHC*, Nucl. Phys. Proc. Suppl. **222-224** (2012) 61–80, arXiv:1201.1295 [hep-ph].

[23] K. Jakobs, *Physics at the LHC – From Standard Model measurements to Searches for New Physics*, in *Proceedings, 2011 European School of High-Energy Physics (ESHEP 2011)*, pp. 309–358. 2014. arXiv:1206.7024 [hep-ex].

[24] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1–91, arXiv:hep-ph/0409313 [hep-ph].

[25] B. Isildak, *Measurement of the differential dijet production cross section in proton-proton collisions at* $\sqrt{s} = 7$ *tev*. PhD thesis, Bogazici U., 2011. arXiv:1308.6064 [hep-ex].

[26] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv:hep-ph/0011363 [hep-ph].

[27] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[28] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[29] G. P. Salam and G. Soyez, *A Practical Seedless Infrared-Safe Cone jet algorithm*, JHEP **05** (2007) 086, arXiv:0704.0292 [hep-ph].

[30] A. Ahriche and S. Nasri, *Neutrino masses, dark matter and baryon asymmetry of the Universe*, J. Phys. Conf. Ser. **593** (2015) no. 1, 012010, arXiv:1409.2999 [hep-ph].

[31] ATLAS, CMS Collaboration, P. de Jong, *Supersymmetry searches at the LHC*, in *Proceedings, 32nd International Symposium on Physics in Collision (PIC 2012)*, pp. 241–254. 2012. arXiv:1211.3887 [hep-ex].

[32] S. Raychaudhuri, *Extra Spacetime Dimensions and the LHC*, in *Physics at the Large Hadron Collider*, A. Datta, B. Mukhopadhyaya, A. Raychaudhuri, A. K. Gupta, C. L. Khetrapal, T. Padmanabhan, and M. Vijayan, eds., pp. 213–239. 2009. http://inspirehep.net/record/1331184/files/chp_10.1007_978-81-8489-295-6_14.pdf.

[33] E. Eichten, K. D. Lane, and M. E. Peskin, *New Tests for Quark and Lepton Substructure*, Phys. Rev. Lett. **50** (1983) 811–814.

[34] P. Chiappetta and M. Perrottet, *Possible bounds on compositeness from inclusive one jet production in large hadron colliders*, Phys. Lett. **B253** (1991) 489–493.

[35] CMS Collaboration, V. Khachatryan et al., *Search for quark contact interactions and extra spatial dimensions using dijet angular distributions in proton proton collisions at* $\sqrt{s} = 8\ TeV$, Phys. Lett. **B746** (2015) 79–99, arXiv:1411.2646 [hep-ex].

[36] L. R. Evans, *LHC Accelerator Physics and Technology Challenges; rev. version*, Tech. Rep. LHC-Project-Report-303. CERN-LHC-Project-Report-303, Apr, 1999. https://cds.cern.ch/record/386693.

[37] F. Fayette, *Strategies for precision measurements of the charge asymmetry of the W boson mass at the LHC within the ATLAS experiment*. PhD thesis, Paris U., VI-VII, 2009. arXiv:0906.4260 [hep-ex].

[38] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[39] ATLAS Collaboration, E. Stanecka, *The ATLAS Inner Detector operation, data quality and tracking performance*, in *Proceedings, 32nd International Symposium on Physics in Collision (PIC 2012)*, pp. 383–388. 2013. arXiv:1303.3630 [physics.ins-det].

[40] F. Hgging, *Upgrades of the ATLAS Pixel Detector*, tech. rep. https://cds.cern.ch/record/1599958. Sep 2013.

[41] ATLAS Collaboration, K. Sliwa, *"ATLAS Overview and Main Results"*, in *Proceedings, International School on High Energy Physics : Workshop on High Energy Physics in the near Future. (LISHEP 2013)*. 2013. arXiv:1305.4551 [hep-ex].

[42] ATLAS Collaboration Collaboration, G. Aad, S. Bentvelsen, G. J. Bobbink, and Bos, *The ATLAS Experiment at the CERN Large Hadron Collider*, J. Instrum. **3** (2008) S08003. 437 p. https://cds.cern.ch/record/1129811. Also published by CERN Geneva in 2010.

[43] ATLAS Collaboration, G. Aad et al., *Commissioning of the ATLAS Muon Spectrometer with Cosmic Rays*, Eur. Phys. J. **C70** (2010) 875–916, arXiv:1006.4384 [physics.ins-det].

[44] C. Gabaldon, *Performance of the ATLAS Trigger System*, tech. rep., Sep, 2011. https://cds.cern.ch/record/1384802.

[45] ATLAS Collaboration, W. Buttinger, *The ATLAS Level-1 Trigger system*, J. Phys. Conf. Ser. **396** (2012) 012010.

[46] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73** (2013) no. 8, 2518, arXiv:1302.4393 [hep-ex].

[47] S. Hche, *Introduction to parton-shower event generators*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014) Boulder, Colorado, June 2-27, 2014*. 2014. arXiv:1411.4085 [hep-ph].

[48] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[49] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[50] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:hep-ph/0206293 [hep-ph].

[51] T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G. P. Salam, F. Siegert, P. Starovoitov, and M. Sutton, *A posteriori inclusion of*

parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project, Eur. Phys. J. **C66** (2010) 503–524, arXiv:0911.2985 [hep-ph].

[52] J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, et al., *CT10 next-to-next-to-leading order global analysis of QCD*, Phys.Rev. **D89** (2014) no. 3, 033009, arXiv:1302.6246 [hep-ph].

[53] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[54] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. **C75** (2015) no. 5, 204, arXiv:1412.3989 [hep-ph].

[55] S. Forte, E. Laenen, P. Nason, and J. Rojo, *Heavy quarks in deep-inelastic scattering*, Nucl. Phys. **B834** (2010) 116–162, arXiv:1001.2312 [hep-ph].

[56] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. **D82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[57] V. Balagura, *Notes on Van der Meer scan for absolute luminosity measurement*, arXiv:1103.1129 [physics.ins-det].

[58] T. Golling, H. S. Hayward, P. U. E. Onyisi, H. J. Stelzer, and P. Waller, *The ATLAS Data Quality Defect Database System*, Eur. Phys. J. **C72** (2012) 1960, arXiv:1110.6119 [physics.ins-det].

[59] ATLAS Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, Tech. Rep. ATLAS-CONF-2013-083, CERN, Geneva. http://cds.cern.ch/record/1570994.

[60] *Data Preparation Check List For Physics Analysis*, Website, 06-Nov-2011. https://twiki.cern.ch/twiki/bin/viewauth/Atlas/DataPreparationCheckListForPhysicsAnalysis.

[61] ATLAS Collaboration Collaboration, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2010-038, CERN, Geneva, Jul, 2010. https://cds.cern.ch/record/1277678.

[62] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Eur. Phys. J. **C 73** (2013) no. 3, 2304, arXiv:1112.6426 [hep-ex].

[63] *Tile Trip Reader*, Website, 25-Sep-2012. https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TileTripReader.

[64] T. Barillari, E. Bergeaas Kuutmann, T. Carli, and Erdmann, *Local Hadronic Calibration*, Tech. Rep. ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001, CERN, Geneva, Jun, 2008.

[65] ATLAS Collaboration, G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) no. 1, 17, arXiv:1406.0076 [hep-ex].

[66] M. Cacciari and G. P. Salam, *Dispelling the* N^3 *myth for the* k_t *jet-finder*, Phys. Lett. **B 641** (2006) 57–61, arXiv:hep-ph/0512210 [hep-ph].

[67] *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, Tech. Rep. ATLAS-CONF-2015-037, CERN, Geneva, Aug, 2015. http://cds.cern.ch/record/2044941.

[68] T. Carli, A. Cooper-Sarker, C. Meyer, V. Kus, B. Malaescu, F. Mueller, S. Shimizu, P. Starovoitov, S. Yanush, and G. Vardanyan, *Measurement of the inclusive jet cross section in proton-proton collision data at 2.76 TeV centre-of-mass energy using the ATLAS detector*, Tech. Rep. ATL-PHYS-INT-2014-011, CERN, Geneva, Jun, 2014. https://cds.cern.ch/record/1706248.

[69] G. D'Agostini, *Improved iterative Bayesian unfolding*, ArXiv e-prints (Oct., 2010) , arXiv:1010.0632 [physics.data-an].

[70] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. **A362** (1995) 487–498.

[71] A. Dattagupta, J. Horton, L. Kogan, S. Lammers, G. Lefebvre, B. Malaescu, A. Olariu, P. Starovoitov, and M. Vetterli, *Determination of the jet energy scale and resolution at ATLAS using Z/γ-jet events in data at $\sqrt{s} = 8$ TeV*, Tech. Rep. ATL-COM-PHYS-2014-791, CERN, Geneva, Jul, 2014. https://cds.cern.ch/record/1741697.

[72] J. Lacey, A. Pilkington, W. Leight, M. Queitsch-Maitland, G. Vardanyan, and F. Ruehr, *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8$ TeV*, Tech. Rep. ATLAS-COM-CONF-2014-068, CERN, Geneva, Aug, 2014. https://cds.cern.ch/record/1751412.

[73] *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at sqrt(s) = 8 TeV*, Tech. Rep.

ATLAS-CONF-2015-002, CERN, Geneva, Mar, 2015.
https://cds.cern.ch/record/2001682.

[74] S. Baker, G. Brown, M. Campanelli, T. Carli, and Cavasinni, *Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s}$=7 TeV using the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2011-738, CERN, Geneva, Jun, 2011. https://cds.cern.ch/record/1360174.

[75] ATLAS Collaboration, *Jet energy resolution and reconstruction efficiencies from in-situ techniqes with the ATLAS Detector using proton-proton collisions at a center-of-mass energy $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-054, 2010.

[76] Baker, S and others, *Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s}$=7 TeV using the ATLAS detector*, ATL-COM-PHYS-2011-738, 2011. http://cdsweb.cern.ch/record/1360174.

[77] N. Boelaert, M. Campanelli, T. Carli, E. Ertel, and Feng, *Measurement of high mass dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector (Supporting Documentation)*, Tech. Rep. ATL-COM-PHYS-2012-1250, CERN, Geneva, Aug, 2012.

[78] G. Bohm, *Introduction to statistics and data analysis for physicists*. Verl. Dt. Elektronen-Synchrotron, Hamburg, 2010.

[79] B. Efron, *An introduction to the bootstrap*. Chapman & Hall, New York, 1994.

[80] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, arXiv:1101.0538 [hep-ph].

[81] *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*, Tech. Rep. ATL-PHYS-PUB-2011-009, CERN, Geneva, Jul, 2011. https://cds.cern.ch/record/1363300.

[82] A. Buckley, D. Kar, P. Loch, and J. Monk, *Summary of ATLAS Pythia 8 tunes*, Tech. Rep. ATL-COM-PHYS-2012-738, CERN, Geneva, Jun, 2012. https://cds.cern.ch/record/1453344.

[83] S. Gieseke, C. Rohr, and A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. **C72** (2012) 2225, arXiv:1206.0041 [hep-ph].

[84] S. Dittmaier, A. Huss, and C. Speckner, *Weak radiative corrections to dijet production at hadron colliders*, JHEP **1211** (2012) 095, arXiv:1210.0438 [hep-ph].

[85] ATLAS Collaboration, G. Aad et al., *Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. **D86** (2012) 014022, arXiv:1112.6297 [hep-ex].

[86] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP **03** (2011) 015, arXiv:1011.2268 [hep-ph].

[87] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, *Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks*, Phys. Rev. Lett. **101** (2008) 142001, arXiv:0806.0848 [hep-ph].

[88] J. Gallicchio, J. Huth, M. Kagan, M. D. Schwartz, K. Black, and B. Tweedie, *Multivariate discrimination and the Higgs + W/Z search*, JHEP **04** (2011) 069, arXiv:1010.3698 [hep-ph].

[89] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. **B659** (2008) 119–126, arXiv:0707.1378 [hep-ph].

[90] J. Gallicchio and M. D. Schwartz, *Pure Samples of Quark and Gluon Jets at the LHC*, JHEP **10** (2011) 103, arXiv:1104.1175 [hep-ph].

[91] ATLAS Collaboration, G. Aad et al., *Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. **C74** (2014) no. 8, 3023, arXiv:1405.6583 [hep-ex].

[92] A. Sherstnev and R. S. Thorne, *Parton Distributions for LO Generators*, Eur. Phys. J. **C 55** (2008) 553–575, arXiv:0711.2473 [hep-ph].

[93] ATLAS Collaboration Collaboration, *Light-quark and Gluon Jets in ATLAS*, Tech. Rep. ATLAS-CONF-2011-053, CERN, Geneva, Apr, 2011. https://cds.cern.ch/record/1342550.

[94] ATLAS Collaboration, G. Aad et al., *Search for Quark Contact Interactions in Dijet Angular Distributions in pp Collisions at $\sqrt{s} = 7$ TeV Measured with the ATLAS Detector*, Phys. Lett. **B694** (2011) 327–345, arXiv:1009.5069 [hep-ex].

[95] J. Gallicchio and M. D. Schwartz, *Quark and Gluon Tagging at the LHC*, Phys. Rev. Lett. **107** (2011) 172001, arXiv:1106.3076 [hep-ph].

[96] LHCb Collaboration, X. Cid Vidal, *Muon Identification in the LHCb experiment*, in *Proceedings, 45th Rencontres de Moriond on Electroweak Interactions and Unified Theories*. 2010. arXiv:1005.2585 [hep-ex]. http://inspirehep.net/record/855624/files/arXiv:1005.2585.pdf.

[97] D. Krohn, M. D. Schwartz, T. Lin, and W. J. Waalewijn, *Jet Charge at the LHC*, Phys. Rev. Lett. **110** (2013) no. 21, 212001, arXiv:1209.2421 [hep-ph].

[98] CMS Collaboration Collaboration, *Performance of quark/gluon discrimination in 8 TeV pp data*, Tech. Rep. CMS-PAS-JME-13-002, CERN, Geneva, 2013. https://cds.cern.ch/record/1599732.

[99] J. Pumplin, *How to tell quark jets from gluon jets*, Phys. Rev. **D44** (1991) 2025–2032.

[100] A. J. Larkoski, G. P. Salam, and J. Thaler, *Energy Correlation Functions for Jet Substructure*, JHEP **1306** (2013) 108, arXiv:1305.0007 [hep-ph].

[101] ATLAS Collaboration, T. A. collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, .

[102] *Testing and Validation*, https://atlas-rtt.cern.ch/prod/docs/userguide/andhttp://atlas-proj-computing-tdr.web.cern.ch/atlas-proj-computing-tdr/Html/Computing-TDR-31.htm.

[103] *Documenting Python*, https://docs.python.org/devguide/documenting.html.

[104] *ATLAS offline software tutorial*, 2009. http://atlas.fis.utfsm.cl/atlas/ATLASSOFT.pdf.

[105] *CMT in ATLAS*, http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/sit/Policy/.

[106] *Data Analysis in ATLAS*, 2013. https://indico.cern.ch/event/217511/contribution/17/attachments/349324/486982/ATLAS_Analysis_Model_2013.03.13.pdf.